

STIC-Biotech/ChemLib

137546

mej

From: Whiteman, Brian
Sent: Wednesday, November 10, 2004 10:02 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

10/069,386, Alaoui-Jamali et al.
2/19/02

search SEQ ID NO: 2 against databases.

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

RECEIVED
NOV 10 2004
TECHNICAL INFORMATION
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search
NA Sequence: # _____
AA Sequence :# _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

RESULT 4
AAG74742
ID AAG74742 standard; protein; 157 AA.
XX AC AAG74742;
XX DT 03-SEP-2001 (first entry)
XX DE Human colon cancer antigen protein SEQ ID NO: 5506.
XX KN Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX KW colorectal carcinoma.
XX OS Homo sapiens.
XX PN WO200122920-A2.
XX PD 05-APR-2001.
XX 28-SEP-2000; 2000WO-US026524.
XX PR 29-SEP-1999; 99US-0157137P.
XX PR 03-NOV-1999; 99US-0163220P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
DR N-PSDB; AAH34147.
XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers.
PS Claim 11; Page 7098-7099; 9803PP; English.
XX AAH322943 to AAH37195 and AAG75154 to AAG7788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytotoxic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent sequences used in the exemplification of the present invention. N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO: 1027 XX SQ Sequence 157 AA.

RESULT 5
AEO60015
ID AEO60075 standard; protein; 142 AA.
XX AC AEO60075;
XX DT 29-JUL-2004 (first entry)
XX DE Human genome derived single exon protein #6309.
XX KW Human; gene expression; single exon probe; microarray; alternative splicing event; genomic alteration.
XX OS Homo sapiens.
XX PN US2003194704-A1.
XX PD 16-OCT-2003.
XX PF 03-APR-2002; 2002US-00009386.
XX PR 03-APR-2002; 2002US-00029386.
XX PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX PI Penn SG, Rank DR, Hanzel DK;
XX DR WPI; 2004-119264/12.
XX PT New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative splicing events, for assessing genomic alterations or as tools for surveying tissues.
XX Claim 45; SEQ ID NO 33709; 80pp; English.

CC The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression comprising a plurality of single exon nucleic acid probes cited above where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality, a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to

a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DOCID=20030194704

Sequence 142 AA:					
Query Match	62.5%	Score	645.5;	DB	8;
Best Local Similarity	96.8%	Pred. No.	7.4e-55;		Length 142;
Matches	122;	Conservative	0;	Mismatches	3;
				Indels	1;
				Gaps	1
71	LAPAPALPPPEFLGDFDSATISILRELDTSMDGTTEPQNPTVPLGLQNEVPPQPD				130
18	LPPLPC-PPEFLGDFDSATISILRELDTSMDGTTEPQNPTVPLGLQNEVPPQPD				76
131	FVFLRAISRYLGDSIDDDFFLDIDTSAVAKEPARAPPEPHNLFCAPGSMEWNLHDHM				190
77	PVFLRAISRYLGDSIDDDFFLDIDTSAVAKEPARAPPEPPENLFCAPGSMEWNLHDHM				136
191	EEILGS				196
137	EEILGS				142

DLT 6
0074

ABO60074 standard; protein; 111 AA.

ABO 60074;

229-JUL-2004 {first entry}

卷之二

THEORY AND PRACTICE IN THE FIELD OF HUMAN RELATIONS

Human; gene expression; single exon alternative splicing event; genomic

Homo sapiens.

USS2003194704-A1.

16-OCT-2003.

03-APPB-2002: 20021115-00029385

卷之三

03-APR-2002; 2002US-00029386.

(PENN /) PENN S G.
(BANK /) BANK S D.

(JUNK / RANK D R.
(HANZ / HANZEL D K.

Penn SG. Rank DB. Hanze] DK.

WBDI: 2004 110264/13

MEI, 2004-13284/12.

New human genome-derived single exon gene expression analysis

splicing events, for assessing genomic

Claim 45: SEQ ID NO. 33708: 8000: English

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are spatially-addressable sets of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records [each record including data on the expression of a single exon probe cited above]. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising gross alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.htm?DocID=200301094704

Sequence 111 AA;

	Query Match	Score	DB 8;	Length	
1	89.9% Locality	402;	DB 8;	111;	
2	90%; Conservative	Pred. No.	3-4e-31;		
3	0;	Mismatches	9;	Indels	0;
4	Gaps	0;			
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					

T 7
004
AAV36004 standard; protein; 236 AA.
AAV36004
13-SEP-1999 (first entry)
Extended human secreted protein sequence, SEQ ID NO. 389.
Secreted protein; human; cytokine; cellular proliferation; cell movement;
cellular differentiation; immune system regulator; anti-inflammatory;
haemopoiesis regulator; tissue growth regulator; tumour inhibitor;
reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
genetic disease.
Homo sapiens.

PD 24-JUN-1999. PD 10-TUN-2004.
 XX XX
 PF 17-DEC-1998; 98WO-IB002122. PF 15-OCT-2001; 2001US-00978360.
 XX XX
 PR 17-DEC-1997; 97US-00699571P. PR 17-DEC-1998; 98WO-IB002122.
 PR 09-FEB-1998; 98US-0074121P. PR 09-FEB-1999; 99WO-IB00282.
 PR 13-APR-1998; 98US-0081563P. PR 21-TUN-2000; 2000WO-IB000951.
 PR 10-AUG-1998; 98US-0096116P. PR 15-SEP-2000; 2000US-00663600.
 XX XX
 (GEST) GENSET SA. (GEST) GENSET SA.
 PA Dumas Milne Edwards J, Bougueret L, Jobert S, Cluseel C;
 XX PI Dumas Milne Edwards J, Bougueret L, Jobert S, Cluseel C;
 XX PI Ducleart A;
 XX DR WPI; 2004-440404/41.
 DR N-PSDB; ADP1807.
 XX PS
 PT New isolated Polynucleotide encoding secreted polypeptide, useful for
 PT Gene therapy, or in diagnostic procedures to identify individuals having
 PT genetic diseases resulting from abnormal expression of the genes.
 XX PS
 PT Claim 2: SEQ ID NO 568; 113PP; English.
 XX
 CC The invention relates to human cDNA sequences that encode human secreted
 CC proteins. The invention also relates to an antibody that specifically
 CC binds to a polypeptide of the invention and a method of binding the
 CC polypeptide to an antibody. The Polynucleotides are useful for expressing
 CC the entire secreted proteins which they encode and for distinguishing
 CC human tissues and cells from non-human tissues and cells, and for
 CC distinguishing between human tissues and cells that do or do not express
 CC the polynucleotides comprising the cDNAs. The polynucleotides and
 CC polypeptides are useful in forensic procedures or diagnostic procedures
 CC to identify individuals with genetic diseases resulting from abnormal
 CC expression of the genes corresponding to the cDNAs. The sequences are
 CC also useful in gene therapy to control or treat genetic diseases. This
 CC sequence represents a human secreted polypeptide of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html.
 XX SQ Sequence 236 AA;
 Query Match 20.9%; Score 216; DB 2; Length 236;
 Best Local Similarity 32.6%; Pred. No. 1.3e-12;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
 SQ Sequence 236 AA;
 Query 4 GLKRKHSDLEEE----ERNEWSPAQGQSYQQA-----LARISDKVQRSLGPRAPS 51
 Db 5 GLKRREBEEKEPLAVDSW-WLDEGHAAYAQAPAVASSSLEPLSVKLHHSQXPD 63
 Query 52 LRRHVLHNFQLQOAAALRPLAPALPPEL-----FIGBEDFLSATIGSIRE 101
 Db 64 LRHLVXNTLIRRQAS--MAAAALPVPPTPAPXAVDNLLASSDAAASMAXILED 121
 Query 102 LDTSMGTEPPQNPVTPGLQNEVPPD----FVFEAL----SSRYLGDQSLDDFL 152
 Db 122 L-SHIEGLSQAPQ----LADEGPGRS1GGXPXLLGPAIGLQDNLQGLEGFL 175
 Query 153 DIDTSAKEKE--PARAPPBP-PHNLPCAGSWE---WNELDHIMEILGS 196
 Db 176 DIDTSMYDNEELWAPASEGKRPGPED--GPGKEEAPELDEAELDYLMDVLVT 225
 Query 153 DIDTSAKEKE--PARAPPBP-PHNLPCAGSWE---WNELDHIMEILGS 196
 Db 176 DIDTSMYDNEELWAPASEGKRPGPED--GPGKEEAPELDEAELDYLMDVLVT 225
 RESULT 8
 ADP19312
 ID ADP19312 standard; protein; 236 AA.
 AC ADP19312;
 XX DT 26-AUG-2004 (first entry)
 XX DB Human secreted polypeptide #163.
 XX AC Human secreted protein; genetic disease.
 XX OS Homo sapiens.
 XX PN US2004110939-A1.
 XX
 RESULT 9
 AAY44362
 ID AAY44362 standard; protein; 236 AA.
 AC AAY44362;
 XX DT 14-MAR-2000 (first entry)
 XX DB Human cell cycle regulation protein-3.

CC CECRP-3; cell cycle regulation protein-3; cell proliferation; cell proliferative disease; cancer; atherosclerosis; cirrhosis; hepatitis; psoriasis; immune system disorder; allergy; asthma; acquired immune deficiency syndrome; Crohn's disease; Blast method; rheumatoid arthritis; gene therapy; chromosomal mapping.

OS Homo sapiens.

XX Key Modified-site 44 Location/Qualifiers
 PT /note= "Potential phosphorylation site"
 GT Modified-site 60
 PT /note= "Potential phosphorylation site"
 GT Modified-site 73
 PT /note= "Potential phosphorylation site"
 GT Modified-site 98
 PT /note= "Potential phosphorylation site"
 GT Modified-site 117
 PT /note= "Potential phosphorylation site"
 GT Modified-site 123
 PT /note= "Potential phosphorylation site"
 GT Modified-site 180
 PN WO994593-A2.
 XX PDD 16-DBC-1999.
 XX PE 08-JUN-1999; 99WO-US012906.
 XX PR 08-JUN-1998; 98US-00886952.
 PA (INCYT-) INCYTE PHARM INC.
 XX PI Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Baughn MR;
 Patterson C;
 DR WPI:2000-105887/09.
 DR N-PSDB; AAZ29482.

PT Novel regulatory proteins, for diagnosis, treatment and prevention of cell proliferative and immune system diseases.
 XX
 BS: Claim 1: Page 70: 88pp; English.
 CC The present sequence is cell cycle regulation protein-3 (CECRP-3). Prints analytical method was used to identify this protein. CECRPs are activators of cell proliferation or inhibitors of cellular processes that modulate proliferation. They are used to treat or prevent cell proliferative diseases like cancers, atherosclerosis, cirrhosis, hepatitis, psoriasis, immune system disorders (e.g. acquired immune deficiency syndrome, allergy, asthma, Crohn's disease, rheumatoid arthritis). Antibodies are raised to screen for specific binding agents. The corresponding nucleic acid is used in gene therapy, chromosomal mapping and isolation of related sequences

XX Sequence 236 AA;

Query Match 20.6%; Score 213; DB 3; Length 236;
 Best Local Similarity 32.6%; Pred. No. 2.6e-12;
 Matches 76; Conservative 36; Mismatches 6; Indels 52; Gaps 13;
 Gaps 13;
 QY 4 GLRKHSDEEE--ERNEWSPAGLQSYQQA-----LRLRSLDKYQRSLGPRAPS 51
 DD 5 GLRKREEEEKPLAVDSW-WLDPGHAAVADAPPAVASSSLFLDSVLKHLHSLOOEPD 63
 QY 52 LRHHLTHNTLQQLQALRLLPAPAPLPEPEPL-----FLGERSDEPSLSATIGSILRE 101
 DD 64 LRHLVLYNTLRLQIAS-MAPAAALPPVSPPAAPSVADNLASSDAASMSLLED 121
 QY 102 LDTSMDGTEPPNPVPTPLGLQNEVPP-----QPDDEVFLAL--SSRYLGDSGLDDFFL 152
 DD 122 L-SHIEGLSQAPP-----LADEGPPGRSGIGAAPSGLGDLGQEGLEGLFE 175
 QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPSWE----WNELDHMEIILGS 196
 DB 176 DDTSMYDNELWAPASEGLKPGPED--GPKEEADELDEAELDYLMDVLTG 225

RESULT 10

AAM93724
 ID AAM93724 standard; protein; 236 AA.
 XX
 AC AAM93724;
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide, SEQ ID NO: 3677
 XX Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX Homo sapiens.
 XX EP1130094-A2.
 XX PD 05-SEP-2001.
 XX PF 07-JUL-2000; 2000EP-00114089.
 XX PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kamai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI: 2001-524255/58.
 DR N-PSDB; AAK94674.

XX 830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation.
 XX
 PS Claim 8; SEQ ID NO 3677; 1380pp + Sequence Listing: English.
 XX The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated CC and nucleotide sequences of 5' - and 3' -ends of the cDNA molecules have CC been determined. Primers for synthesising the full length cDNA are useful CC for clarifying the function of the protein encoded by the cDNA. The full CC length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo capping method. The primers CC enable the production of the full length cDNA easily without any special CC methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention. Note: The sequence data for this patent did CC not form part of the printed specification, but was obtained in CD-ROM format directly from EPO
 XX Sequence 236 AA;

Query Match 20.6%; Score 213; DB 4; Length 236;
 Best Local Similarity 32.6%; Pred. No. 2.6e-12;
 Matches 76; Conservative 36; Mismatches 6; Indels 52; Gaps 13;
 Gaps 13;
 QY 4 GLRKHSDEEE--ERNEWSPAGLQSYQQA-----LRLRSLDKYQRSLGPRAPS 51
 DD 5 GLRKREEEEKPLAVDSW-WLDPGHAAVADAPPAVASSSLFLDSVLKHLHSLOOEPD 63
 QY 52 LRHHLTHNTLQQLQALRLLPAPAPLPEPEPL-----FLGERSDEPSLSATIGSILRE 101
 DD 64 LRHLVLYNTLRLQIAS-MAPAAALPPVSPPAAPSVADNLASSDAASMSLLED 121
 QY 102 LDTSMDGTEPPNPVPTPLGLQNEVPP-----QPDDEVFLAL--SSRYLGDSGLDDFFL 152
 DD 122 L-SHIEGLSQAPP-----LADEGPPGRSGIGAAPSGLGDLGQEGLEGLFE 175

Qy 153 DIDTSAVEKE--PARAPEP-PHNLFCAPGSWE---WNELDHIMEIILGS 196
 Db 176 DIDTSMYDNEIWLAPASEGKPSPED---GPCKEEAPEBLDEAELDYMIVGT 225

RESULT 011
 ID AAG89292 standard; protein, 236 AA.
 XX
 DT 11-SEP-2001 (first entry)
 DE Human secreted protein, SEQ ID NO: 412.
 XX
 KW GENSET.
 XX
 OS Homo sapiens.
 PN WO200142451 A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 07-DEC-2000; 2000WO-IB0001938.
 XX
 PR 08-DEC-1999; 99US-0169620P.
 PR 06-MAR-2000; 2000US-0187470P.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Bougueret L, Jobert S;
 DR WPI; 2001-367870/38.
 N-PSDB; AAH64895.
 XX
 PT Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of PT diseases, and for diagnosis of those diseases.
 XX
 PS Claim 21; Page 889-890; 921pp; English.
 XX
 CC The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cell's own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptides expression and activity. The present sequence is a GENSET polypeptide of the invention.

XX
 SQ Sequence 236 AA;

Query Match 20.6%; Score 213; DB 4; Length 236;
 Best Local Similarity 32.6%; Pred. No. 2.6e-12;
 Matches 36; Mismatches 69; Indels 52; Gaps 13;

Qy 4 GLKRKHSPLZEER---ERWENSPAGLQSYKQAA----LLRISLDKVRQSLGPRAPS 51
 5 GLKRKREBEEKEPLAEDSW WLDPGHTAVQAPPAVASSSLFDISVLKHHSP1QSQED 63

Db 52 LRRHVLIHTLQOQLQOALRLAPAPALPPEPL----FLGEEBDFSLSATIGSILRE 101

Db 64 LRHLVLTWNTLRLRQAS--MAPAAALPPVPSPPAAPSVDNLASSDALSASMASLIED 121
 Qy 102 LDTSMGTEPPQNVPQLQNEVP-----QDPVPLAEAL--SSRYLGPSGLDDFL 152
 Db 122 L-SHIEGLSQAPQP---LADEEPGRSIGGAAPSLGALDLPATCCLLDGEGLIFE 175
 Qy 153 DIDTSAVEKE--PARAPEP-PHNLFCAPGSWE---WNELDHIMEIILGS 196
 Db 176 DIDTSMYDNEIWLAPASEGKPSPED---GPCKEEAPEBLDEAELDYMIVGT 225

RESULT 12
 ID ADL31644 standard; protein, 236 AA.
 XX
 AC ADL31644;
 XX
 DT 20-MAY-2004 (first entry)
 DE Human protein encoded by a full length cDNA clone SeqID 3677.
 XX
 KW human; medicine; signal transduction; glycoprotein; transcription; KW oligo-capping method.
 XX
 Homo sapiens.
 XX
 PN EP1396543-A2.
 XX
 PD 10-MAR-2004.
 XX
 PF 07-JUL-2000; 20003EP-00025638.
 XX
 PR 08-JUL-1999; 99JP-00104486.
 PR 11-JAN-2000; 2000JP-001183865.
 PR 02-MAY-2000; 2000JP-00114089.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Otsuki T, Koga H;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 WPI; 2004-204755/20.
 DR N-PSDB; ADL31643.

XX
 PT New Oligonucleotide primers (830 cDNAs) useful for synthesizing full length human cDNAs.

XX
 PS Example 1; SEQ ID NO 3677; 1340pp; English.

XX
 CC This invention relates to a novel primers useful for synthesising full length cDNA molecules that encode human proteins. Specifically, it refers to secretory or membrane proteins that are potential therapeutic agents/ target molecules in the field of medicine, and in particular genes/ encoding proteins that are associated with signal transduction, glycoproteins and transcription. The present invention describes a method for efficiently cloning a full length human cDNA from both the 5' and 3' ends using the oligo-capping method. This polypeptide sequence is a full length human protein of the invention.

XX
 SQ Sequence 236 AA;

Query Match 20.6%; Score 213; DB 8; Length 236;
 Best Local Similarity 32.6%; Pred. No. 2.6e-12;
 Matches 36; Mismatches 69; Indels 52; Gaps 13;

Qy 4 GLKRKHSPLZEER---ERWENSPAGLQSYKQAA----LLRISLDKVRQSLGPRAPS 51
 5 GLKRKREBEEKEPLAEDSW WLDPGHTAVQAPPAVASSSLFDISVLKHHSP1QSQED 63

Db 52 LRRHVLIHTLQOQLQOALRLAPAPALPPEPL----FLGEEBDFSLSATIGSILRE 101

Qy

XX 18-MAR-1999; 99WO-US005939.
 XX PD 16-OCT-2003.
 PR 20-MAR-1998; 98US-0078803P.
 PR 03-APR-2002; 2002US-00029386.
 XX PR 03-APR-2002; 2002US-00029386.
 PA (GEMY) GENETICS INST INC.
 XX PA (PENN/) PENN S G.
 PI Jacobs K, MCCOY JM, Lavallie ER, Collins-Racie LA, Evans C;
 PR (RANK/) RANK D R.
 PA (HANZL/) HANZL D K.
 XX PI Penn SG, Rank DR, Hanzel DK;
 DR WPI; 1999-562059/47.
 XX PI XX DR WPI; 2004-11964/12.
 PR New polynucleotides derived from murine fetal cell cDNA libraries, potentially used as, e.g. vaccines.
 XX PS Claim 13(a); Page 94; 107PP; English.
 XX PS This is the predicted amino acid sequence of a novel human secreted protein, g21.1, as deduced from an isolated adult blood cDNA clone (see ANZ19894). The invention provides new human secreted proteins (see ANZ1822-38) and polynucleotides (see AAZ19893-901) isolated from foetal cell, adult blood, adult brain and foetal kidney cDNA libraries. They are predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data are given. Suggested activities include nutritional, cytokine, tissue growth, cell proliferation and differentiation, immunostimulant (e.g. as vaccine), immunosuppressive, haemopoiesis, regulating, activin or inhibin, chemotactic or chemokinetic, haemostatic or thrombolytic, receptor/ligand activity, antiinflammatory, cathearin or tumour invasion suppressor, and tumour inhibition activities.
 XX SQ Sequence 236 AA;
 Query Match 20.4%; Score 211; DB 2; Length 236;
 Best Local Similarity 32.6%; Pred. No. 4.1e-12;
 Matches 76; Conservative 35; Mismatches 70; Indels 52; Gaps 13;
 QY 4 GLKKEKHSIDIEEE---ERWEWSFAGLQSYQQC-----LIRISIDKVQRSLGPRAPS 51
 DB 5 GLKKEKREBEKEPLAVDSW-WLDPGHAAVAVQAPPVAVSSLFDSLWKLHESLQOSEPD 63
 QY 52 LRRHVLHINTLQLQAAALRAPAPALPPEL-----F1GEEDEFSLSATIGSTIRE 101
 DB 64 LRHLVLYNTLRRQAS--MAPAALPPVPSPPAAPSADNLASSMASLIED 121
 QY 102 LDTSMDGTTEPPQNPTVPLQIQLQNEVPP-----QDPDFLEAL-SSRYLGDSGLDDFL 152
 DB 122 L-SHIEGLSQAPQ-----LAXEPPGRSIIIGGAAPSGLQADLGPATCCLJDGLECFE 175
 QY 153 DIDTSAVEKE---PARAPEP-PENNLFAPGSNE---WNELDHMEILGS 196
 DB 176 DIDTSMYDNLWAPASEGLXPKPGED--GPGKSEAPELDEAELDLYMDLVGT 225
 XX SQ Sequence 222 AA;
 Query Match 19.8%; Score 205; DB 8; Length 222;
 Best Local Similarity 32.6%; Pred. No. 1.5e-11;
 Matches 75; Conservative 34; Mismatches 69; Indels 52; Gaps 13;
 QY 4 GLKRKHSIDLEBEE---ERNEWSPAGLQSYQQA-----LRLRSLDKVRQSLQPRAPS 51
 DB 5 GLKRKREEEEBEEKEPLAVDSW-WLDDQHTAYAQAPPVAVSSLFDSLWKLHESLQOSEPD 63
 QY 52 LRRHVLHINTLQLQAAALRAPAPALPPEL-----FLGBEDFSLSATIGSILRE 101
 DB 64 LRHLVLYNTLRRQAS--MAPAALPPVPSPPAAPSADNLASSMASLIED 121
 QY 102 LDTSMDGTTEPPQNPTVPLQIQLQNEVPP-----QDPDFLEAL-SSRYLGDSGLDDFL 152
 AB060367 standard; protein; 222 AA.
 AB060367 Human genome derived single exon protein #6601.
 AB060367 Human; gene expression; single exon probe; microarray;
 DT 29-JUL-2004 (first entry) alternative splicing event; genomic alteration.
 XX Homo sapiens.
 XX US2003194704-A1.
 PN

Mon Nov 15 17:33:06 2004

142 L-SHFGLSQAPOP----LADEGPGRSGGAAPSGLGDLGCGCGLFE 175
153 DIDTSAYKE - PARASPEP-PHNIFCAIGSWE----WNLGDHMEII 193
176 DIDTSMYDNEWLWAPASEGILKPGED---GZKEZAPELDEEELDYLNDVL 222

Search completed: November 15, 2004, 14:03:14
Dog time : 159 secs

Scoring table:	BLOSUM62	ALIGMENTS
Gapopen:	10.0	Gapext: 0.5
Searched:	1825181 seqs, 575374646 residues	RESULT 1
Total number of hits satisfying chosen parameters:	1825181	Q9UJW9
Minimum DB seq length:	0	ID: Q9UJW9
Maximum DB seq length:	2000000000	PRELIMINARY:
Post-processing:	Minimum Match 0%	AC: Q9UJW9,
	Maximum Match 100%	DT: 01-MAY-2000 (TRMBLrel. 13, Created)
Database :	UniProt 02:*	DT: 01-MAY-2000 (TRMBLrel. 13, Last sequence update)
	1: uniprot_sprot:*	DT: 01-MAR-2004 (TRMBLrel. 26, Last annotation update)
	2: uniprot_trembl:*	DE: RPA-binding trans-activator.
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		GN: Name=RAT1;
		OS: Homo sapiens (Human)
		OC: Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
		OX: NCBI_TaxID:9606;
		RN: [1]
		RP: SEQUENCE FROM N.A.
		RX: MEDLINE=20403390; PubMed=10982866;
		RA: Cho J.M., Song D.J., Aloui-Jamali M.A.;
		RT: "RBT1, a novel transcriptional co-activator, binds the second subunit of Replication Protein A";
		RL: Nucleic Acids Res. 28:3418-3485 (2000).
		DR: EMBL; AF192529; AAF05761.1; -.
		DR: InterPro; IPR002263; SERTA.
		DR: PFam; PF06031; SERTA; 1.
		SEQUENCE 196 AA; 21798 MW; 33433EF7F8A9BEA7 CRC64;
		Query Match 100.0%; Score 1033; DB 2; Length 196;
		Best Local Similarity 100.0%; Pred. No. 1.1e-72; Mismatches 0; Indels 0; Gaps 0;
		Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
		QY: 1 MEGGLKEKHSDSLREEERWERWSQAPGQSYQQALLRISLDKVRQSRGAPSLLRRLVLIHN 60
		Db: 1 MEGGLKEKHSDSLREEERWERWSQAPGQSYQQALLRISLDKVRQSRGAPSLLRRLVLIHN 60
		QY: 61 TQQQQLAQLRAPAPALPPEFLGEGDFESLATIQSILRLLDTMDGTTEBQNPYPLIG 120
		Db: 61 TQQQQLAQLRAPAPALPPEFLGEGDFESLATIQSILRLLDTMDGTTEBQNPYPLIG 120
		QY: 121 LQEVPPDPPVFLAEALSRVLYGSDGLDDFLIDDSAVEKEPARAPPEPHNLQAPGS 180
		Db: 121 LQEVPPDPPVFLAEALSRVLYGSDGLDDFLIDDSAVEKEPARAPPEPHNLQAPGS 180
		RESULT 2
		Q96CQ2
		ID: Q96CQ2
		PRELIMINARY:
		AC: Q96CQ2,
		DT: 01-DBC-2001 (TRMBLrel. 19, Created)
		DT: 01-DBC-2001 (TRMBLrel. 19, Last sequence update)
		DE: RPA-binding trans-activator.

DE	Name=SERPARD3 ; Homo sapiens (Human) ; Chordata; Craniata; Vertebrata; Euteleostomi ; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo ; NCBI_TaxID=9606;	Replication protein-binding trans-activator RBT1 (SERTA domain containing 3).
GN		Name=SERPARD3 ; Mus musculus (Mouse) ;
OS		Eukaryota; Metazoa; Chordata; Rodentia; Muridae; Murinae; Mus ;
OC		Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
OC		[1]
RN	SEQUENCE FROM N.A.	
RC	TISSUE=Ovary, and Uterus;	
RX	MEDLINE=22388257; PubMed=12477932;	
RA	Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G., Wagner L., Schaefer C.F., Bhat N.K., Altshuler S.F., Zeeberg B., Buetow K.H., Schuler G.D., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Stapleton M., Soares M.B., Bonaldo M.F., Rubin G.M., Hong L., Brownstein M.J., Usdin T.B., Toshyuki S., Carninci P., Prange C., Brownstein M.J., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., Loquellano N.A., Peters J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Tochman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grinwood J., Schmitz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schein J.E., Jones S.J., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences," Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
RA	SEQUENCE FROM N.A.	
RC	TISSUE=Ovary;	
RA	Strausberg R.; Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.	
RA	SEQUENCE FROM N.A.	
RC	TISSUE=Uterus;	
RA	Strausberg R.; Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: BC14061; AAH14061.1; -.	
DR	EMBL: BC50643; AAH10643.1; -.	
DR	InterPro: IPR009263; SERTA.	
DR	Pfam: PF06031; SERTA.	
DR	SEQUENCE 196 AA; 21768 MW; 9C54AFA126F17AF1 CRC64;	
DR	Query Match 99.3%; Score 1026; DB 2; Length 196; Best Local Similarity 99.3%; Pred. No. 3.9e-72; Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
DR	1 MEGGLKRSPLLEEEBERWNSPAGLQSYQALLRISLDKVRSLSGPRASLRRHLVHN 60	
DR	1 MVGGGLKRSPLLEEEBERWNSPAGLQSYQALLRISLDKVRSLSGPRASLRRHLVHN 60	
DR	61 TLQQIQAALRAPAPALPAPPFLGEGEDFSLSATIGSILRQLDTSMDGTBPPQNPVTPLG 120	
DR	61 TLQQIQAALRAPAPALPAPPFLGEGEDFSLSATIGSILRQLDTSMDGTBPPQNPVTPLG 120	
DR	61 TLQQLQALRAPAPALPAPPFLGEGEDFSLSATIGSILRQLDTSMDGTBPPQNPVTPLG 120	
DR	121 LQNEYPPQPPDFVFLAASSRLGDSGLDDFLDIDTSAVKEPARAPPPEPHNLFCAPGS 180	
DR	121 LQNEYPPQPPDFVFLAASSRLGDSGLDDFLDIDTSAVKEPARAPPPEPHNLFCAPGS 180	
DR	181 WENNELDHINBILGS 196	
DR	181 WENNELDHINBILGS 196	
RESULT 3	Q9ERC3	PRELIMINARY;
AC	Q9ERC3	PRT; 197 AA.
DR	01-MAR-2001 (TREMBLrel. 16, last sequence update)	RESULT 4
DR	01-MAR-2001 (TREMBLrel. 16, last annotation update)	STDL_MOUSE
DR	01-OCT-2004 (TREMBLrel. 28, last annotation update)	ID_STDL_MOUSE STANDARD;
DR	182 WENNELDHINBILGS 197	PRT; 236 AA.

AC Q9JL10; Q925E6; Q9D888;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE SERTA domain-containing protein 1 (Transcriptional regulator interacting with the PHD-bromodomain 1) (TRIP-B1) (CDK4-binding protein p34SEB1) (SEB1-1).
 DE Name-Sertadl; Synonyms=seel1;
 OS Mus musculus (Mouse);
 OC Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus;
 OX NCBI_TAXID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA "Cloning of mouse SEB1-1 cDNA";
 RL Submitted (DDBJ-1999) to the EMBL/GenBank/DBJU databases.
 RN [2]
 RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTIONS.
 MEDLINE=2:23:1173; PubMed=113:1592;
 RX Hsu S.-I., Yang C.M., Sim K.G., Hentschel D.M., O'Leary E.,
 RA Bonventre J.V.;
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamamoto J., Miyosawa H.,
 RT "TRIP-Br: a novel family of PHD zinc finger- and bromodomain-
 interacting proteins that regulate the transcriptional activity of
 E2F-1/DP-1";
 RL EMBO J. 20:2273-2285 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Embryo, and Small intestine;
 RX MEDLINE=B2:23:5463; PubMed=101266;
 RA Okazaki Y., Furuno M., Kasukawa T., Asachi J., Bono H., Kondo S.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldi R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradic D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.W.,
 RA Gaasterland T., Pontius J.U., Qi D., Ramachandran S.,
 RA Grimond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D.,
 RA Kana A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongaya A., Kurochkin I.V., Lee H., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Mischoni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavani W.J., Peretea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Reeder D.J., Reed J., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Secou M., Shimada K.,
 RA Sultana R., Takekawa Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahrlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yamagisawa M., Yang L.,
 RA Hirono K., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Shiraki T., Wakita K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Miyazaki A., Sakai K., Sasaki D., Shishiba K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs,"
 RL Nature 422:563-573 (2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 MEDLINE=22:38:257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klaauwer R.D., Collins F.S., Wagner L., Sherman C.M., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Hopkins R.F., Jordan H., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Diatchenko L., Squires M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Stapleton M., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Hara S.S., Loqueland J., Peters R.D., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., Loqueland J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., McElwain P.J., McKernan K.J., Hale S., Garcia A.M., Gay J.J., Hulyk S.W.,
 RA Villalobos D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinwood J., Schmitz J., Myers R.M.,
 RA Buttermfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marrs M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Acts at E2F responsive promoters to integrate signals
 CC provided by PHD- and/or bromodomain-containing transcriptional activity. Renders
 CC factors Stimulates E2F-1/DP-1 transcriptional activity. Renders
 CC the activity of cyclin D1/CDK4 resistant to the inhibitory effects
 CC of p16INK4a.
 CC -!- SUBUNIT: Interacts with the PHD-bromodomain of TRIP1, TRIM28/TIF1B
 CC and P300/CBP. Binds to DP1. Also interacts with CDK4.
 CC -!- SIMILARITY: Belongs to the TRIP-Br Family.
 CC -!- SIMILARITY: Contains 1 SERTA domain.
 CC -!- SIMILARITY: Contains 1 SERTA domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -!- SIMILARITY: Contains 1 SERTA domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -!- SIMILARITY: Contains 1 SERTA domain.

DR AF21291; AAF21653.1; -;
 DR AF36400; AAK52829.1; -;
 DR AF36401; AAK52830.1; -;
 DR AF36402; BAB2130.1; -;
 DR AF008303; BAB2558.1; -;
 DR BC16077; AAH16077.1; -;
 DR MGII:19:13438; Sertad1.
 DR InterPro: IPR002263; SERTA.
 DR Pfam: PF06031; SERTA; 1.
 DR Transcription regulation.
 KW SERTA.
 DR DOMAIN 45 82
 FT CONFLICT 9 10 KR -> NG (in Ref. 3; BAB25588).
 FT CONFLICT 12 12 E -> K (in Ref. 2; AAK52829).
 SQ SEQUENCE 236 AA; 25136 MW; 989ADF8299DE84C5 CRC64;
 Query Match 21.0%; Score 217; DB 1; Length 236;
 Best Local Similarity 31.5%; Pred. No. 5e-09; RA
 Matches 76; Conservative 33; Mismatches 64; Indels 68; Gaps 12;
 RA 4 GLERKHSDDLEEEERERNW-----W-----SPAGLOS-----YQQALRISLDKVRQSGLGP 48
 RA 5 GLKRK-----REBETMEALSVDSDWCLDPSPHPAVAVPTVASSFLDFLSVVKLHSRLRQS 60
 RA 49 APSLRHVLHINTLQLQALR LAPAPALPPEPL-----FGEEDFSLSATIGSI 98
 RA 61 EPLRHLVWVTRRQASHE-----PAPVLPPEPIQPAPSPVADSLASSDAGLSMASI 118
 RA 99 LRELDTSMDGTGPQNPVTPLGLQNYVPP-----QDPDVFLAE-----SSRYLGDGGLDD 149
 RA 119 LEPNHEEDLNQAPQPA-----DGPSPGSIGSPNGLDQPLGATGCLDDLEG 172
 RA 150 FFLDIDTSAVKE----PAR-----APPEPHNLFCAPGSWEWNLHDHMTTIG 195
 RA 173 LFDDIDTSMYDSELWLDASEGILKPGPENGPAKEEPP-----ELDEAEDLYMDVLVG 224
 DR 196 S 196
 DR 225 T 225

RESULT 5
 STD1_HUMAN
 ID STD1_HUMAN
 AC O9HBV2; O9HBV7;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)

CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AP17959; AAFA0849.1; -
DR	EMBL; AP366402; AAK52831.1; -
DR	EMBL; AY130850; AAM77300.1; -
DR	EMBL; BC02670; AAH02610.1; -
DR	Genew; HGNC:17932; SERTD1.
DR	GO; GO:0008284; P: Positive regulation of cell proliferation; TAS.
DR	GO; GO:000079; P: Regulation of CDK activity; TAS.
DR	InterPro; IPR009263; SERTA.
DR	Pfam; PF06031; SERTA; 1.
KW	Polymorphism; Transcription regulation.
FT	SENTA.
FT	A -> T (in dbSNP:268687).
FT	/FTId=VAR015581.
SEQ	SEQUENCE 236 AA; 24673 MW; 92F25EA328P155B1 CRC64;
Query Match	20.6%; Score 213; DB 1; Length 236;
Best Local Similarity	32.6%; Prid. No. 1e-08;
Matches	76; Conservative 36; Mismatches 69; Indels 52; Gaps 137;
QY	4 GLRKHSIDLEEE---ERWEWSPGAGLQSYQQA-----LRLISLDKVRQSLGPPRAPS 51
Db	5 GLRKREEEEERERPLAIDSW-WLDPQHAAAVAAQAPPAYASSIFDLSVKLHSLQSEPD 63
QY	52 LRRHVLHINTLQLQAAIRLAPALPHEPL-----FLGEEDFELSATIGSILRE 101
Db	64 LRRHVLVVTNLRRQAS-MADAAFLPPVPSPPAAPSVDNLASSDAALASMASLLED 121
QY	102 LDTSMDGTEPPQNPVPLGLQNEVP-----QPDVFLERL 152
Db	122 L-SHIEGLSQAPQ-----LADGPGRSIGAAPSGLDILGPATGCLDDGLFBE 175
QY	153 DITDSAVEKE---PARAPPEP_PHLFCAGPSWE---WNELDHIMEIILGS 196
Db	176 DITDSMYDNELWAPASESLKPGFED---GPKRKEAPPELDEAFLDYLMDVLVGT 225
RESULT 6	-----
Q6P771	PRELIMINARY;
ID	Q6P771
AC	Q6P771;
DT	05-JUL-2004 (TREMBLrel. 27, Created)
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE	Hypotheoretical protein.
OS	Rattus norvegicus (Rat).
OC	Eutherota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.
NCBI_TaxID	10116;
[1]	SEQUENCE FROM N.A., AND VARIANT THR-31.
REDLINE#22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;	-----
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	STRASBURG R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klauser R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,	Klauser R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
Altenschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	Altenschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Hsieh F.,	Hopkins R.F., Jordan H., Moore T., Max S.I., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnaratre P.H.W.,	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnaratre P.H.W.,
Richards S., Madan A., Young A.C., Shevchenko Y., Sanchez A.,	Richards S., Madan A., Young A.C., Shevchenko Y., Sanchez A.,
Vahey J., Helton E., Ketteman M., Madan A., Rodriguez S.,	Vahey J., Helton E., Ketteman M., Madan A., Rodriguez S.,
Whiting M., Madan A., Touchman J.W., Schmutz J., Bouffard G.G.,	Whiting M., Madan A., Touchman J.W., Schmutz J., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Schmutz J., Bouffard G.G.,	Blakesley R.W., Touchman J.W., Schmutz J., Bouffard G.G.,
Rodriguez A.C., Grimwood J., Myers R.M., Garcia A.M., Gay L.J., Hulyk S.W.,	Rodriguez A.C., Grimwood J., Myers R.M., Garcia A.M., Gay L.J., Hulyk S.W.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
"Generation and initial analysis of more than 15,000 full-length human	"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";	and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
-1- FUNCTION: Acts at E2F responsive promoters to integrate signals	-1- FUNCTION: Acts at E2F responsive promoters to integrate signals
provided by PHD- and/or bromodomain-containing transcription	provided by PHD- and/or bromodomain-containing transcription
factors. Stimulates E2F1/DP-1 transcriptional activity. Renders	factors. Stimulates E2F1/DP-1 transcriptional activity. Renders
the activity of cyclin D1/CDK4 resistant to the inhibitory effects	the activity of cyclin D1/CDK4 resistant to the inhibitory effects
of p16 (INK4a).	of p16 (INK4a).
-1- SUBUNIT: Interacts with the PHD-bromodomain of TIF1, TRIM28/TIF1B	-1- SUBUNIT: Interacts with the PHD-bromodomain of TIF1, TRIM28/TIF1B
and P300/CBP. Binds to DPL. Also interacts with CDK4.	and P300/CBP. Binds to DPL. Also interacts with CDK4.
-1- SIMILARITY: Belongs to the TRIP-BP family.	-1- SIMILARITY: Belongs to the TRIP-BP family.
-1- SIMILARITY: Contains 1 SERTA domain.	-1- SIMILARITY: Contains 1 SERTA domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration	This SWISS-PROT entry is copyright. It is produced through a collaboration
between the European Bioinformatics Institute and the EMBL outstation -	between the European Bioinformatics Institute and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its	the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way	use by non-profit institutions as long as its content is in no way
misleading or misleading.	misleading or misleading.

RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR009233; SERTA.

Pfam: PF06031; SERTA; 1.

KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 236 AA; 25419 MW;

88B7502E0473EDBB CRC64;

Query Match 20.1%; Score 208; DB 2; Length 236;

Best Local Similarity 31.9%; Pred. No. 2.5-08;

Matches 76; Conservative 33; Mismatches 67; Indels 62; Gaps 13;

Qy 4 GLRKHSIDLEEEERWE----W-----SPGLOSYQQLAIRLAPALPPPEL-----FGEEDFESIATIG 96

Db 5 GLRK----REBEEETMESLSDVSDWLDQSRPVAQTPTAVAS--SSLFDSVVKLHSSLR 58

Qy 47 PRAPLRLRIVLHNTLQQQALAIRLAPALPPPEL-----FGEEDFESIATIG 96

Db 59 QSEPDRLHLYLVNTLRRQASME--PTTVPPEPIQTAPSVDNFLESSDAGLSASMA 116

Qy 97 SILEPLDTSMGTEPPONPVTPGLQNEVPP-----QDPVPFLEAL--SSRYLGDSL 147

Db 117 SILEPLSHIEDLNQVPPQDA-----DEGPPGRSVGGVLPNLGDLQPATGCLLDDGL 170

Qy 148 DDFFLDIDTSAVEKE--PARAPPEP-PHNLFCAQSWEW----NEDLHMEILIGS 196

Db 171 EGLFEDDTSMDSLWLPASEGLRKPVEN--GPAKEETPELDEAEIDLMDVLVGT 225

RESULT 8

Q9DCZ2 PRELIMINARY; PRT: 244 AA.

AC Q9DCZ2; PRELIMINARY; PRT: 244 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2004 (TREMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Mus musculus adult male kidney cDNA, RIKEN full-length enriched

DE Library, clone:0610008D10 product:TRIP-BRI, full insert sequence.

DN Name=sertad1; ID Q9DCZ2;

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN 11

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RX MEDLINE=99279255; PubMed=10346636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RT Meth. Enzymol. 303:19-44 (1999).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RX MEDLINE=2108566; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RN [3]

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RA The RIKEN Genome Exploration Research Group Phase I & II Team;

RA The RIKEN Genome Consortium,

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RT Nature 420:1685-690 (2001).

RN [4]

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shikata Y., Hayatsu N., Sugahara Y., Shidata K., Itoh M.,

RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RT Genome Res. 10:1617-1630 (2000).

RN [5]

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shikata Y., Hayatsu N., Sugahara Y., Shidata K., Itoh M.,

RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RT Genome Res. 10:1617-1630 (2000).

RN [5]

RT and mouse cDNA sequences.";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR009233; SERTA.

Pfam: PF06031; SERTA; 1.

KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 236 AA; 25419 MW;

88B7502E0473EDBB CRC64;

Query Match 20.1%; Score 208; DB 2; Length 236;

Best Local Similarity 31.9%; Pred. No. 2.5-08;

Matches 76; Conservative 33; Mismatches 67; Indels 62; Gaps 13;

Qy 4 GLRKHSIDLEEEERWE----W-----SPGLOSYQQLAIRLAPALPPPEL-----FGEEDFESIATIG 96

Db 5 GLRK----REBEEETMESLSDVSDWLDQSRPVAQTPTAVAS--SSLFDSVVKLHSSLR 58

Qy 47 PRAPLRLRIVLHNTLQQQALAIRLAPALPPPEL-----FGEEDFESIATIG 96

Db 59 QSEPDRLHLYLVNTLRRQASME--PTTVPPEPIQTAPSVDNFLESSDAGLSASMA 116

Qy 97 SILEPLDTSMGTEPPONPVTPGLQNEVPP-----QDPVPFLEAL--SSRYLGDSL 147

Db 117 SILEPLSHIEDLNQVPPQDA-----DEGPPGRSVGGVLPNLGDLQPATGCLLDDGL 170

RESULT 8

Q9DCZ2 PRELIMINARY; PRT: 244 AA.

AC Q9DCZ2; PRELIMINARY; PRT: 244 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2004 (TREMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Mus musculus adult male kidney cDNA, RIKEN full-length enriched

DE Library, clone:0610008D10 product:TRIP-BRI, full insert sequence.

DN Name=sertad1; ID Q9DCZ2;

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN 11

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RX MEDLINE=99279255; PubMed=10346636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RT Meth. Enzymol. 303:19-44 (1999).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RX MEDLINE=2108566; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RN [3]

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RA The RIKEN Genome Exploration Research Group Phase I & II Team;

RA The RIKEN Genome Consortium,

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RT Nature 420:1685-690 (2001).

RN [4]

SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=Kidney;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shikata Y., Hayatsu N., Sugahara Y., Shidata K., Itoh M.,

RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RT Genome Res. 10:1617-1630 (2000).

RN [5]

Name=Cdc44; Synonyms=HEPP;
Mus musculus ('Mouse').
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
NCBI_TaxID=10090; [1]

SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Bone marrow;
MEDLINE=23234683; PubMed=14646851; DOI=10.1038/nature01266;
Okazaki Y., Furukawa M., Kasukawa T., Adachi J., Bono H., Kondo S.,
Nikaido I., Osato N., Saito R., Suzuki H., Yamakata I., Kiyosawa H.,
Yagi K., Tonaru Y., Hasegawa Y., Nogami T., Schonback C., Gojobori T.,
Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
Schrimal L.M., Karapin A., Matsuda H., Batalov S., Beisel K.W.,
Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
Grimmond S., Gusinovich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
Kanai A., Kawaji H., Kawasaki Y., Kezierski R.M., King B.L.,
Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
Maglaj D.R., Malta L., Marchionni L., McKenzie L., Miki H., Pesole G.,
Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y.,
Wilmung L.G., Wyrshik Boris A., Yanagisawa M., Yang I., Yang L.,
Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
Hirozaki N., Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
Hara A., Hashizume W., Imotani K., Itohi Y., Itoh M., Kawara I.,
Miyazaki A., Sakurai K., Sasaki D., Shibata K., Shinagawa A.,
Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
Birney E., Hayashizaki Y.,
"Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.", Nature 420:563-573 (2002). [2]

SEQUENCE OF 13-237 FROM N.A.
TISSUE=Breast tumor;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2422603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmier C.M., Schulz G.D.,
Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Dianchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Buetow K.H., Schaefer C.F., Blatt N.K.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarstam P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalona D.K., Munoz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bonifard G.G.,
Bucklesley R.W., Toucman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield S.N., Krzywinski M.I., Skalak D.E.,
Schnerge A., Schein J.E., Jones S.J.M., Marra M.A.,
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:16893-16903 (2002). [3]

- FUNCTION: May be involved in molecular regulation of hematopoietic
stem cells and progenitor cell lineage commitment and
differentiation.

-!- TISSUE SPECIFICITY: Expressed preferentially in hematopoietic progenitors and mature blood cells. Expressed at low levels in the heart, lung, spleen, and thymus and at a higher level in muscle.

-!- DEVELOPMENTAL STAGE: Developmentally regulated. Preferential expression in both fetal and adult hematopoietic progenitors and mature blood cells during embryonic and adult hematopoiesis.

-!- SIMILARITY: Belongs to the TRIP-Br family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AX010531; BAB27012; -.

DR EMBL; AX322228; AAK31074; 1; -.

DR EMBL; AX03980; BRC28119; 1; -.

DR EMBL; BCO12953; AAH12953; 1; ALT_INIT.

DR MGD:MGI:1919213; Cdca4.

DR InterPro; IPRO09263; SERTA_Pfam; PP06031; SERTA; 1.

DR CONFLICT; 196 196 S -> T (in Ref. 3).

FT CONFLICT 213 213 T -> A (in Ref. 3).

FT SEQUENCE 237 AA; 26107 MW; 52EC046B2DDF326B CRC64; -.

Query Match 13.2%; Score: 136; DB: 1; Length: 237;

Best Local Similarity 26.1%; Pred. No. 0.01;

Matches 63; Conservative 32; Mismatches 86; Indels 60; Gaps 9;

97

Qy 4 GLKRKHDLEEEPERWENSPA-GLOQSYQALLRISLDKVKQNSLGPRAFSLRRHVLIIHNTL 62

Db 5 GLKRKYGDDEBEGVFGFTVPSVSLQ--RQSLLDMSVLPQLQCHMLVVEENLCSVLIANTV 62

Qy 63 QQIQLAQLR -----LAP -APALPPEPLFLGEBDSLUSATIGSILRLEDTSMDGTTEPPQ 113

Db 63 RQIQCQEMSDQGVHMGMAQNVDPAPVERLVSLE - -ILCRTVRGAEEBHPAPELDAPLQ 119

Qy 114 NPVTPGLQNEVPPQDP-----VFEALSSRYLGDPSGLDDEFF 151

Db 120 NSVSEBLPIPGSAPQRNPQNSLPOSSILWMDSPQENRNSFQKSLLDQIETLENK - -NSSSVBLF 177

Qy 152 LDIDTSAVEKEPAR-----APPEPPPHNLFCAPGSNEWNLDHIMBII 193

Db 178 SDVDSYYSDLDTIVTGMNSGTKSLLCNGLEGFAAATPPSSTKS - -DLAELDHVVBL 234

Qy 194 L 194

Db 235 V 235

RESULT 12 AAH55824 PRELIMINARY; PRT; 237 AA.

RP SEQUENCE FROM N.A.

TISSUE_Eye; PRELIMINARY; PRT; 237 AA.

RX MEDLINE:22389257; PubMed:12479392;

RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Bhat N.K., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Wang J., Hsieh P., Hopkins R.F., Jordan H., Moore T., Max S.I., Marusika K., Farmer A.A., Rubin G.M., Hong L., Risch N.

RX NCBI_TaxID:10090;

RN [1]

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustini B., Toshiyuki S., Carninci P., Prange C., RA Raha S.S., Loqueilano N.A., Peters G.J., Abramson R.D., Mullany S.J., RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H., RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., RA Villalobos D.K., Mizny D.M., Soderstrom E.J., Lu X., Gibbs R.A., RA Fahey J., Helton E., Petteman M., Madan A., Rodrigues S., Sanchez A., RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., RA Krzywinski M.I., Skalska U., Smialius D.E., Schnarch A., Schein J.E., RA Jones S.J., Marrs M.A., "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences," Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RP	SEQUENCE FROM N.A.	DE (Neuronal/epithelial high affinity glutamate transporter, system Xag), DE member 1 (SLC1A1).
RC	TISSUE=Kidney;	GN Name=C101N13.8;
RX	MEDLINE=22388257; PubMed=122477932;	OS Brachydanio rerio (zebrafish) (Danio rerio)
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	OC Lukayota; Metzao; Chordata; Cratata; Vertebrata; Butelostomi;
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,	OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
RA	Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,	OC Cyprinidae; Danio.
RA	Diachenko L., Matrosina K., Farmer A.A., Rubin G.M., Hong L.,	NCBI_TAXID=7955;
RA	Stashev M., Soares M.F., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	[1]
RA	Brownstein M.J., Usoskin T.B., Toshioyuki S., Carrinici P., Prange C.,	SEQUENCE FROM N.A.
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaiay S.J.,	RA Kimberley A;
RA	Bosak S.A., McIwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	RL Submitted (JUN 2003) to the EMBL/GenBank/DBJ databases.
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulkyk S.W.,	DR EMBL: AL772154; CAD1094.1; -.
RA	Villionon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	DR InterPro: IPR09263; SERTA.
RA	Fahey J., Heilton E., Kettman M., Madan A., Rodotin Y., Sanchez A.,	DR Pfam: PF06031; SERIA, 1.
RA	Whiting M., Madan A., Young A.C., Bouffard G.G.,	SQ SEQUENCE 383 AA; 40402 MW; 76514201CDCF92F CRC64;
RA	Blakesley R.W., Tinchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grinwood J., Schmitz J., Myers R.M., Butterfield Y.S.,	Query Match 12.9%; Score 133; DB 2; Length 393;
RA	Krzyniowski M.I., Skalska S., Smialski D.E., Schnirer A., Schein J.B.,	Best Local Similarity 22.3%; Pred. No. 0.031; Gaps 17;
RA	Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full length human	Matches 81; Conservative 35; Mismatches 71; Indels 176; Gaps 17;
RT	and mouse cDNA sequences.;"	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Kidney;	
RA	Strasberg R.	
RL	Submitted (MAR 2004) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; BC07135; AAH7135.1; -.	
KW	Hypothetical protein	
SO	SEQUENCE 361 AA; 37995 MW; 59CE36B09843BF CRC64;	
Query Match 12.9%; Score 133; DB 2; Length 361;		
Best Local Similarity 22.3%; Pred. No. 0.029; Gaps 17;		
Matches 81; Conservative 35; Mismatches 71; Indels 176; Gaps 17;		
QY	4 GLKRKHSIDLBEERERNEW----SPAGLQ----SY----QALLRISLDKVQSLGPAPS 51	Db 144 SQSPALSASSSLTSPSSGLNSNPAPLACLTPLASLLEBDVNSLCTSPLAPPAPSRLSPS 203
DB	5 GAKRK--LDDEDEGEGKALAAAGGAEGLSKVSYTLOQTIFNMSLMKLYNHRATVPS 61	Db 145 QNPVTPPL 119
QY	52 LRRHVLJHNTLQQLQALR-----	Db 146 TSMGDTTEPP-----
DB	62 LEKRVLJINNMRLRRIQDEKLQEGNRLPFLPPSPBPPDPYDESTREPOAFVLSMVAFF 121	Db 147 VARDSSSALEDEEECPSPLEPTAASAGATSPSPPLQLCPSLNSGALDSKDCSKPCSK 263
QY	72 APAPALP-----PEPL-----FGEEDFSLSA-----	Db 148 94 ----TIGSILRELD-----TSMGDTTEPP-----
DB	122 SQSPALSASSSLTSPSSGLNSNPAPLACLTPLASLLEBDVNSLCTSPLAPPAPSRLSPS 181	Db 149 204 VARDSSSALEDEEECPSPLEPTAASAGATSPSPPLQLCPSLNSGALDSKDCSKPCSK 263
QY	94 ----TIGSILRELD-----TSMGDTTEPP-----	Db 150 120 --GL----QNEVPPQPD----PVFLB----ALSRSYLGDSGLDD-FFLDDTSAVEK 161
DB	182 VARDSSSALEDEEECPSPLEPTAASAGATSPSPPLQLCPSLNSGALDSKDCSKPCSK 241	Db 151 264 LEGLVPLAERSAVNPNTETPLPNNSLDMSTSPPASSSGFLTDLADLDFADDTSMDF 323
QY	120 --GL----QNEVPPQPD----PVFLB----ALSRSYLGDSGLDD-FFLDDTSAVEK 161	Db 152 162 EP-----ARAPPEPPHNLFCAPGSNEWNLDHIMEVLT 193
DB	242 DEGLVPLAERSAVNPNTETPLPNNSLDMSTSPPASSSGFLTDLADLDFADDTSMDF 301	Db 153 324 DPCTSSSGAAPSKLAPMVTADELLKTFPSYSGAAPAVSSN---QPFKMOLTELDHIMEVLT 380
QY	162 EP-----	Db 154 QY 194 LGS 196
DB	302 DPCTSSSGAAPSKLAPMVTADELLKTFPSYSGAAPAVSSN---QPFKMOLTELDHIMEVLT 358	Db 155 361 VGS 363
QY	194 LGS 196	
DB	359 VGS 361	
RESULT 15		
SD7Z227	PRELIMINARY;	PRT; 383 AA.
AC Q7Z227;		
DT 01-JUN-2003 (TREMBLrel. 24, Created)		
DE 01-JUN-2003 (TREMBLrel. 24, Last sequence update)		
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE SI:C101N13.8 (Novel protein similar to human scutle carrier family 1		

Copyright (c) 1993 - 2004 CompuGen Ltd.

DOM protein - protein search, using SW model

Run on: November 15, 2004, 14:06:47 ; Search time 142 Seconds
(without alignments)
488.370 Million cell updates/sec

Title: US-10-069-386-2

Perfect score: 1033

Sequence: 1 MEGGIKRKHSDEEEERWE. APGSNEWNLDHMMIILGS 196

Scoring table: BLOSUM62

Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Published Applications AA: *
1: /cgn2_6/protdata/2/pubdata/jus0_pubcom_pub: *
2: /com2_6/protdata/2/pubdata/jus0_pubcom_pub: *

Database : Sequence 4456, Ap
Sequence 4, Appli
Sequence 2083, Ap
Sequence 3214, Ap
Sequence 536, Ap
Sequence 195088, Ap
Sequence 64002, Ap
Sequence 1813, Ap
Sequence 188300, Ap
Sequence 8565, Ap
Sequence 137248, Ap
Sequence 3, Appli
Sequence 3472, Ap
Sequence 312, Appli
Sequence 480, Ap
Sequence 81, Appli
Sequence 82, Appli
Sequence 54984, Ap
Sequence 6, Appli
Sequence 166325, Ap
Sequence 56782, Ap
Sequence 221982, Ap
Sequence 187376, Ap
Sequence 257033, Ap
Sequence 16, Appli
Sequence 1, Appli
Sequence 58, Appli
Sequence 58, Appli
Sequence 58, Appli
Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-106-698-5516 Application US/10106698
 / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

1: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

10: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

11: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

12: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

13: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

14: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

15: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

16: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

17: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

18: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

19: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

20: / Sequence 5516, Application US/10106698
 / Publication No. US2003109690A1
 / GENERAL INFORMATION:
 / APPLICANT: Ruben et al.
 / TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Peptides
 / FILE REFERENCE: PA005PL

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB ID	Description
1	791	76.6	157	US-10-106-698-5516	Sequence 5516, App
2	645	62.5	142	US-10-029-386-33709	Sequence 33709, A
3	402	38.9	111	US-10-029-386-33708	Sequence 33708, A
4	216	20.9	236	US-09-978-360A-568	Sequence 568, App
5	213	20.6	236	US-09-978-362-412	Sequence 412, App
6	213	20.6	236	US-09-876-997-412	Sequence 412, App
7	213	20.6	278	US-10-296-11-1005	Sequence 1005, App
8	205	19.8	222	US-10-029-386-34001	Sequence 34001, A
9	136	13.2	237	US-10-076-069-2	Sequence 2, Appli
10	125	12.1	314	US-10-170-325-31	Sequence 31, Appli
11	125	12.1	314	US-10-75-889-627	Sequence 627, App
12	106	10.3	254	US-09-925-302-596	Sequence 596, App
12	106	10.2	254	US-09-925-302-596	Sequence 596, App

Query Match Score 791; DB 14; Length 157;
 Best Local Similarity 98.0%; Pred. No. 1.3e-63;
 Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0

Ov 44 SLGRAPSLRHLVHTLLOOALPAPALPAPPPLFGEEDFSLSATIGSLRELD 103

卷之三

5 ALGPRAPSRLRHLVHLNTLQLQALRLPAPALPPEPFLGEEDFSLSAIGSILRELD 64
 Db. 104 TSMGDTEPPONPVTPLGLQNEVPDPVLEALSSRYLGDGIDDFIDTDSAVEKEP 163
 Qy. 165 TSMGDTEPPONPVTPLGLQNEVPDPVLEALSSRYLGDGIDDFIDTDSAVEKEP 124
 Db. 164 ARAPPPEPHNLFCAPGSWENELDHIMEILIGS 196
 Qy. 125 ARAPPPEPHNLFCAPGSWENELDHIMEILIGS 157

RESULT 2
 US-10-029-386-13709
 Sequence 33709, Application US/10029386
 Publication No. US20030194704A1

GENERAL INFORMATION:
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
 FILE REFERENCE: AECOMICA-X-2
 CURRENT APPLICATION NUMBER: US/10/029,386
 NUMBER OF SEQ ID NOS: 34288
 SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 33709
 LENGTH: 142
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC010271.5
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
 US-10-029-386-33709

Query Match 62.5%; Score 645.5; DB 14; Length 142;
 Best Local Similarity 96.8%; Pred. No. 1.7e-50; Indels 1; Gaps 1;
 Matches 122; Conservative 3; Mismatches 0; Nucleotides 97;
 Qy. 71 LAPALAPPEPFLGEEDFSLSATIGSILRELDTSMDGHEPPONPVTPLGLQNEVPQPD 130
 Db. 18 LPPLPC-PPPEPFLGEEDFSLSATIGSILRELDTSMDGHEPPONPVTPLGLQNEVPQPD 76
 Qy. 131 PVFLEALSSRYLGDGIDDFIDTDSAVEKEPARAPPPEPHNLFCAPGSWENELDHIM 190
 Db. 77 PVFLEALSSRYLGDGIDDFIDTDSAVEKEPARAPPPEPHNLFCAPGSWENELDHIM 136
 Qy. 191 EILIGS 196
 Db. 137 EILIGS 142

RESULT 3
 US-10-029-386-33708
 Sequence 33708, Application US/10029386
 Publication No. US20030194704A1

GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 FILE REFERENCE: AECOMICA-X-2
 CURRENT APPLICATION NUMBER: US/10/029,386
 NUMBER OF SEQ ID NOS: 34288
 SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 33708
 LENGTH: 142
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC010271.5
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
 US-10-029-386-33708

Query Match 38.9%; Score 402; DB 14; Length 111;
 Best Local Similarity 89.9%; Pred. No. 1.4e-18; Indels 0; Gaps 9;
 Matches 80; Conservative 0; Mismatches 0; Nucleotides 97;
 Qy. 1 MEGGLRKHSLEEEFRERWSPAGQSQOALLRISLDPGRPLRHLVHLHN 60
 Db. 3 MGGGRKHSLEEEFRERWSPAGQSQOALLRISLDPGRPLRHLVHLHN 62
 Qy. 61 TLQQLQANILAPAPALPPLFGEEDF 89
 Db. 63 TLQQLQANILAPAPALPPLFGEEDF 91

RESULT 4
 US-09-978-360A-568
 Sequence 568, Application US/09978360A.
 Publication No. US2004011039A1

GENERAL INFORMATION:
 APPLICANT: Edwards, Jean-Baptiste Dumas Milne
 APPLICANT: Duclert, Aymeric
 APPLICANT: Bouquelert, Lydie
 APPLICANT: Jobert, Severin
 APPLICANT: Clusel, Catherine
 TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
 FILE REFERENCE: 56-US4-CIP
 CURRENT APPLICATION NUMBER: US/09/978,360A
 CURRENT FILING DATE: 2001-10-15
 PRIORITY NUMBER: US 60/066,677
 PRIOR FILING DATE: 1997-11-13
 PRIORITY NUMBER: US 60/069,957
 PRIOR FILING DATE: 1997-12-17
 PRIORITY NUMBER: US 60/074,121
 PRIOR FILING DATE: 1998-02-09
 PRIORITY NUMBER: US 60/081,563
 PRIOR FILING DATE: 1998-04-13
 PRIORITY NUMBER: US 60/096,116
 PRIOR FILING DATE: 1998-08-10
 PRIORITY NUMBER: US 60/099,273
 PRIOR FILING DATE: 09-04
 PRIORITY NUMBER: US 09/191,997
 PRIOR FILING DATE: 1998-11-13
 PRIORITY NUMBER: US 09/215,435
 PRIOR FILING DATE: 1998-12-17
 PRIORITY NUMBER: PCT/IB98/02122
 PRIOR FILING DATE: 1998-12-17
 PRIORITY NUMBER: US 09/247,155
 PRIOR FILING DATE: 1999-02-09
 PRIORITY NUMBER: US 09/247,155
 NUMBER OF SEQ ID NOS: 810
 SOFTWARE: Patent.PM
 SEQ ID NO: 568
 LENGTH: 236
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SIGNAL
 LOCATION: -31...-1
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (28, 30, 40, 67, 86, 117, 120)
 OTHER INFORMATION: unknown
 US-09-978-360A-568

Query Match 20.9%; Score 216; DB 11; Length 236;
 Best Local Similarity 32.6%; Pred. No. 2.4e-11;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
 Qy. 4 GLKRKHSLEEEFRERWSPAGQSQYQA-----LURISLDRVQRSIGPRAPS 51

5 GLRKREBEEKEPLAVDSW-WLDPGHAAVAQAPPVASSSLFDLSVILHHSLOQXSPD 63
 52 LRRHVLHINTLQLQAAIRLAPAPLPEPL-----FGEEDDSLSATIGSILRE 101
 64 LRHLVLVNTLRLRQAS-MAPAAALPVPPTPAPXVADNLASSMASMAXLED 121
 Db 102 LDTSMDGTEPPQNPTVPLGLQNEVPPD----PFLD-----SSRYLGDSGLDDFL 152
 Qy 122 L-SHIEGLSQARQP----LADEGPPGSSIGGXGXPLXLDLGLPAGTCLLNGLEGIFE 175
 Db 153 DIDTSAVEKE---PARAPPB-PHNLFCAGSWE----WNLDEHMETILGS 196
 Qy 176 DIDTSMYDNLWAPASEGLKPGPD---GP GKKEAPELDEAELDYLMDVLVGT 225
 Db 176 DIDTSMYDNLWAPASEGLKPGPD---GP GKKEAPELDEAELDYLMDVLVGT 225

RESULT 5
 US-09-731-872-412
 Sequence 412, Application US/09731872
 Patent No. US20000102604A1
 GENERAL INFORMATION:
 APPLICANT: Dumas Milne Edwards, Jean Baptiste
 APPLICANT: Bougueret, Lydie
 APPLICANT: Jobert, Severin
 TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
 FILE REFERENCE: US3.REF
 CURRENT APPLICATION NUMBER: US/09/731,872
 CURRENT FILING DATE: 2000-12-07
 PRIOR APPLICATION NUMBER: US 60/169,629
 PRIOR FILING DATE: 1999-12-08
 PRIOR APPLICATION NUMBER: US 60/187,470
 PRIOR FILING DATE: 2000-03-06
 NUMBER OF SEQ ID NOS: 482
 SEQ ID NO: 412
 LENGTH: 236
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-731-872-412

Query Match 20.6%; Score 213; DB 9; Length 236;
 Best Local Similarity 32.6%; Pred. No. 4.6e-11;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

Qy 4 GLRKREBEEKEPLAVDSW-WLDPGHAAVAQAPPVASSSLFDLSVILHHSLOQXSPD 51
 5 GLRKREBEEKEPLAVDSW-WLDPGHAAVAQAPPVASSSLFDLSVILHHSLOQXSPD 63
 Db 52 LRRHVLHINTLQLQAAIRLAPAPLPEPL-----FGEEDDSLSATIGSILRE 101
 64 LRHLVLVNTLRLRQAS-MAPAAALPVPPTPAPXVADNLASSMASMAXLED 121
 Db 102 LDTSMDGTEPPQNPTVPLGLQNEVPPD----PFLD-----SSRYLGDSGLDDFL 152
 Qy 122 L-SHIEGLSQARQP----LADEGPPGSSIGGXGXPLXLDLGLPAGTCLLNGLEGIFE 175
 Db 153 DIDTSAVEKE---PARAPPB-PHNLFCAGSWE----WNLDEHMETILGS 196
 Db 176 DIDTSMYDNLWAPASEGLKPGPD---GP GKKEAPELDEAELDYLMDVLVGT 225

RESULT 7
 US-10-296-115-1065
 Sequence 1065, Application US/10296115
 Publication No. US20040053248A1
 GENERAL INFORMATION:
 APPLICANT: Hyseq Inc
 TITLE OF INVENTION: US 20040053248A1 Nucleic Acids and Polypeptides
 FILE REFERENCE: 784PCT
 CURRENT APPLICATION NUMBER: US/10/296,115
 CURRENT FILING DATE: 2002-11-18
 PRIOR APPLICATION NUMBER: US09/488,725
 PRIOR FILING DATE: 2000-01-21
 PRIOR APPLICATION NUMBER: US09/552,317
 PRIOR FILING DATE: 2000-04-25
 SEQ ID NO: 1478
 NUMBER OF SEQ ID NOS: 1478
 LENGTH: 278
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-296-115-1065

Query Match 20.6%; Score 213; DB 15; Length 278;
 Best Local Similarity 32.6%; Pred. No. 5.6e-11;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

Qy 4 GLRKREBEEKEPLAVDSW-WLDPGHAAVAQAPPVASSSLFDLSVILHHSLOQXSPD 51
 5 GLRKREBEEKEPLAVDSW-WLDPGHAAVAQAPPVASSSLFDLSVILHHSLOQXSPD 63
 Db 52 LRRHVLHINTLQLQAAIRLAPAPLPEPL-----FGEEDDSLSATIGSILRE 101
 64 LRHLVLVNTLRLRQAS-MAPAAALPVPPTPAPXVADNLASSMASMAXLED 121
 Db 102 LDTSMDGTEPPQNPTVPLGLQNEVPPD----PFLD-----SSRYLGDSGLDDFL 152
 Qy 122 L-SHIEGLSQARQP----LADEGPPGSSIGGXGXPLXLDLGLPAGTCLLNGLEGIFE 175
 Db 153 DIDTSAVEKE---PARAPPB-PHNLFCAGSWE----WNLDEHMETILGS 196
 Db 176 DIDTSMYDNLWAPASEGLKPGPD---GP GKKEAPELDEAELDYLMDVLVGT 225

RESULT 6
 US-09-876-97412
 Sequence 412, Application US/09876997
 Publication No. US20030152921A1
 GENERAL INFORMATION:
 APPLICANT: Bougueret, Lydie
 APPLICANT: Jobert, Severin
 APPLICANT: Dumas Milne Edwards, Jean Baptiste
 TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
 FILE REFERENCE: 78 US4 CIP
 CURRENT APPLICATION NUMBER: US/09/876,997
 CURRENT FILING DATE: 2001-06-08
 PRIOR APPLICATION NUMBER: US 09/731,872

218 DIDTSMYDNEWLWAPASEGLKPGPED---GPCKEEAEPLEDAELDYLMDLVGT 267

RESULT 8

US-10-029-386-34001

Sequence 34001, Application US/10029386
Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

SEQ ID NO: 34001

LENGTH: 222

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: MAP TO AC010271.5

OTHER INFORMATION: EXPRESSED IN HEK293, SIGNAL = 0 46

OTHER INFORMATION: SWISSPROT HIT: Q14140, EVALU 2.70e-01

US-10-029-386-34001

Query Match

Score 205; DB 14; Length 222;

Best Local Similarity 32.6%; Pred. No. 2.2e-10;

Matches 75; Conservative 34; Mismatches 69; Indels 52; Gaps 13;

Qy 4 GLRKHSKDSLEEE---ERWNEPAGLQSYQQA---LRLISLDKVQRSILGPAPS 51

Db 5 GLRKREEEEKEPLADSRW-LWDQGHTAVAAPAVASSSLDLSVHLHSLQOSEP 63

Qy 52 LRRHLVHLNQLQAAALRIPAPALPPSL---FGEEDFSLSATIGSLR 101

Db 64 LRHLVLVNVLRRIGAS-MAAAALPPSPPPAAPSVDNLASSAALASMASILED 121

Qy 102 LDTSMDGTPEQNPVTPGLQNEYPP-----OPDPVPEAL--SSRLGDSGLDDFF 152

Db 122 L-SHTEGLSQAQP-----LADEGPPGRIGGAPSLQALDLGSPATGCLLQGLEP 175

Qy 153 DIDTSMYDNEWLWAPASEGLKPGPED---GPCKEEAEPLEDAELDYLMDLV 193

Db 176 DIDTSMYDNEWLWAPASEGLKPGPED---GPCKEEAEPLEDAELDYLMDLV 222

SEQ ID NO: 31

LENGTH: 314

TYPE: PRT

ORGANISM: Homo sapiens

US-10-170-385-31

RESULT 9

Sequence 2, Application US/10076069

Publication No. US2003017714A1

GENERAL INFORMATION:

APPLICANT: NACHTMAN, RONALD

APPLICANT: JURECIC, ROLAND

TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELO

FILE REFERENCE: 39552-176599

CURRENT APPLICATION NUMBER: US/10/076,069

CURRENT FILING DATE: 2003-02-15

PRIOR APPLICATION NUMBER: US 60/268,923

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.0

SEQ ID NO: 2

LENGTH: 317

TYPE: PRT

ORGANISM: Mus musculus

US-10-076-069-2

Query Match

13.2%; Score 136; DB 13; Length 237;

Best Local Similarity 26.1%; Prod. No. 0.00043; Mismatches 86; Gaps 9;

Matches 63; Conservative 32; Mismatches 86; Gaps 9;

Qy 4 GLRKHSKDSLEEE---ERWNEPAGLQSYQQA---LRLISLDKVQRSILGPAPS 62
Db 5 GLRKREEEEKEPLADSRW-LWDQGHTAVAAPAVASSSLDLSVHLHSLQOSEP 62

Qy 63 QOLQANIR-----APAPLEPLIGEEDFSLSATIGSLR 113
Db 63 RQIEMNSQDGWYHGRNAPQNYDRAPEVLYSTE---ILCITVRGAEEHNPAPEDAPLQ 119

Qy 114 NWTTPGLQNEVPPQDP-----VLEAASSRVLGDGLDDFF 151
Db 120 NSVSELPVGSAPGQNPQSLWMDSPQENRGSKBSQSLQIETLENK-NSSSVEUL 177

Qy 152 LDIDTSAVEKEPAR-----APBEPHNLFCAGPSWNEELDHIMEILI 193
Db 178 SDVDSYYDLDVTLVMSGTKSSLNGLEGFAAATPPSSCTKS--DLAELDHWVEIL 234

Qy 194 L 194
Db 235 V 235

RESULT 10

US-10-170-385-31

Sequence 31, Application US/10/170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Raynor, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingman, Susan Mary
; APPLICANT: Krieg, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532652000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 31
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-170-385-31

Query Match 12.1%; Score 125; DB 14; Length 314;
Best Local Similarity 22.7%; Prod. No. 0.006; 88; Indels 120; Gaps 12;
Matches 71; Conservative 34; Mismatches 88; Indels 120; Gaps 12;
Matches 71; Conservative 34; Mismatches 88; Indels 120; Gaps 12;

Qy 2 EGGLKHSKDSLEEEERWNEPAGLQSYQQA---YQQALRISLDKVRSILGPAPSRRH 55
Db 4 KGG-KRKFDSEHDEGLGEK1VSPCDGSKVSYLQRQTINISLMKLYNHRPLTEPSLQKT 62

Qy 56 VLIHTLQQLQAAAR-----PALPPE--PLFGEEDFSLSATIGSLR 103
Db 63 VLNINMRLRQEEELRQEGLSLRPMFTPSOTTERSDYREAPPASFHLSAPSSHPCDLGS 122-
70

Qy 71 -----LAPA-----PALPPE--PLFGEEDFSLSATIGSLR 103
Db 123 TTPLAACLTPASLLEDDDTFTCSQMQTAPKLSPALLPKID-SFSALADEIECLP 181

Qy 104 TSMD-----GTEPPNPVTPGLQNEYVPPQ---PDVPLFLEA-----SSRYJGDS 145
Db 182 TSTSTEAATAATDSVKGTSSEAGTQKLGPQESRADDSKLMDSLPGNFEITSTGFLTDL 241

58 IANTLQQLQAL----RLAP--APALPPEPLFLGEDEFSLSATIGSILRELDTSMDG 108
 75 IANTVRIQIPEEMTQDGTVRVAQAAERAPXDRVLSTE-----ILCRAAWGQEG 123
 109 TEP----PQNPTV----PLGLQNEVPPQDDP-----VFILEAL 138
 124 AHPAPGLGDTGPTVSDLCPTVSAQAPHLLOSSAWMDPRENRGSFKSLDQIFITLE 183
 139 SRYLGDSGLDDEFLDIDTSAVEKEP----AR-----APPEPHNLFCAFGS 180
 184 TK--NPSCKMEELEFSVDSPYDVTGMMGARPGPDEGLSGLAPTPGPSSCKS- 239
 181 WENELDHEMIEL 194
 240 -DGEGLDHWEEILV 252

RESULT 14
 US-10-106-698-4456 Score: 106; DB 13; Length: 241;
 Sequence 4456, Application US/10106698
 Publication No. US20030109690A1

GENERAL INFORMATION: FILE REFERENCE: PA005P1

APPLICANT: Ruben et al.

TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide

CURRENT FILING DATE: 2002-03-17

PRIOR APPLICATION NUMBER: PCT/US03/265524

PRIOR FILING DATE: 2000-09-28

PRIOR FILING DATE: 1999-09-29

PRIOR FILING DATE: 1999-11-03

NUMBER OF SEQ ID NOS: 8564

SEQUENCE: Patent In Ver. 3.0

SEQ ID NO: 4456

LENGTH: 254

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: NAME/KEY: MISC_FEATURE

LOCATION: (105)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-106-698-4456

Query Match Score: 106; DB 14; Length: 254;

Best Local Similarity: 25.6%; Pred. No. 0; 24;

Matches 32; Mismatches 75; Indels 82; Gaps 13;

Db 18 GLRKRCVGHEEDV--EGGALGKTVSSSLQOSILLMSVAKLQMLVEPNLCSVL 74

Qy 58 IANTLQQLQAL----RLAP--APALPPEPLFLGEDEFSLSATIGSILRELDTSMDG 108

Db 75 IANTVRIQIPEEMTQDGTVRVAQAAERAPXDRVLSTE-----ILCRAAWGQEG 123

Qy 109 TEP----PQNPTV----PLGLQNEVPPQDDP-----VFILEAL 138

Db 124 AHPAPGLGDTGPTVSDLCPTVSAQAPHLLOSSAWMDPRENRGSFKSLDQIFITLE 183

Qy 139 SRYLGDSGLDDEFLDIDTSAVEKEP----AR-----APPEPHNLFCAFGS 180

Db 184 TK--NPSCKMEELEFSVDSPYDVTGMMGARPGPDEGLSGLAPTPGPSSCKS- 239

Qy 181 WENELDHEMIEL 194

Db 240 -DGEGLDHWEEILV 252

Query Match Score: 105; DB 13; Length: 241;
 Best Local Similarity: 25.9%; Pred. No. 0; 28;
 Matches 31; Mismatches 79; Indels 76; Gaps 13;

Db 5 GLRKRCVGHEEDV--EGGALGKTVSSSLQOSILLMSVAKLQMLVEPNLCSVL 61

Qy 58 IANTLQQLQAL----RLAPAPALPPEPLFLGEDEFSLSATIGSILRELDTSMDG 111

Db 62 IANTVRIQIPEEMTQDGTVRVAQAAERAPL----DRLVST--EILCRAANGQEGAH 113

Qy 112 -----PQNPTV----PLGLQNEVPPQDDP-----VFILEAL 141

Db 114 APGLGDTGPTVSDLCPTVSAQAPHLLOSSAWMDPRENRGSFKSLDQIFITLE 172

Qy 142 LGDSGLDDPFIDTSAVEKEP-----AR-----APPEPHNLFCAFGS 183

Db 173 -NPSCKMEELEFSVDSPYDVTGMMGARPGPDEGLSGLAPTPGPSSCKS- -DL 228

Qy 184 NELDHMEIIL 194

Db 229 GELDHMEIILV 239

Query Match Score: 105; DB 13; Length: 241;
 Best Local Similarity: 25.9%; Pred. No. 0; 28;
 Matches 31; Mismatches 79; Indels 76; Gaps 13;

Db 4 GLRKRCVGHEEDV--EGGALGKTVSSSLQOSILLMSVAKLQMLVEPNLCSVL 57

Qy 58 IANTLQQLQAL----RLAP--APALPPEPLFLGEDEFSLSATIGSILRELDTSMDG 108

Db 75 IANTVRIQIPEEMTQDGTVRVAQAAERAPXDRVLSTE-----ILCRAAWGQEG 123

Qy 109 TEP----PQNPTV----PLGLQNEVPPQDDP-----VFILEAL 138

Db 124 AHPAPGLGDTGPTVSDLCPTVSAQAPHLLOSSAWMDPRENRGSFKSLDQIFITLE 183

Qy 139 SRYLGDSGLDDEFLDIDTSAVEKEP----AR-----APPEPHNLFCAFGS 180

Db 184 TK--NPSCKMEELEFSVDSPYDVTGMMGARPGPDEGLSGLAPTPGPSSCKS- 239

Qy 181 WENELDHEMIEL 194

Db 240 -DGEGLDHWEEILV 252

Query Match Score: 105; DB 13; Length: 241;
 Best Local Similarity: 25.9%; Pred. No. 0; 28;
 Matches 31; Mismatches 79; Indels 76; Gaps 13;

Db 5 GLRKRCVGHEEDV--EGGALGKTVSSSLQOSILLMSVAKLQMLVEPNLCSVL 61

Qy 58 IANTLQQLQAL----RLAPAPALPPEPLFLGEDEFSLSATIGSILRELDTSMDG 111

Db 62 IANTVRIQIPEEMTQDGTVRVAQAAERAPL----DRLVST--EILCRAANGQEGAH 113

Qy 112 -----PQNPTV----PLGLQNEVPPQDDP-----VFILEAL 141

Db 114 APGLGDTGPTVSDLCPTVSAQAPHLLOSSAWMDPRENRGSFKSLDQIFITLE 172

Qy 142 LGDSGLDDPFIDTSAVEKEP-----AR-----APPEPHNLFCAFGS 183

Db 173 -NPSCKMEELEFSVDSPYDVTGMMGARPGPDEGLSGLAPTPGPSSCKS- -DL 228

Qy 184 NELDHMEIIL 194

Db 229 GELDHMEIILV 239

Query Match Score: 105; DB 13; Length: 241;
 Best Local Similarity: 25.9%; Pred. No. 0; 28;
 Matches 31; Mismatches 79; Indels 76; Gaps 13;

Db 4 GLRKRCVGHEEDV--EGGALGKTVSSSLQOSILLMSVAKLQMLVEPNLCSVL 57

Qy 58 IANTLQQLQAL----RLAP--APALPPEPLFLGEDEFSLSATIGSILRELDTSMDG 108

Db 75 IANTVRIQIPEEMTQDGTVRVAQAAERAPXDRVLSTE-----ILCRAAWGQEG 123

Qy 109 TEP----PQNPTV----PLGLQNEVPPQDDP-----VFILEAL 138

Db 124 AHPAPGLGDTGPTVSDLCPTVSAQAPHLLOSSAWMDPRENRGSFKSLDQIFITLE 183

Qy 139 SRYLGDSGLDDEFLDIDTSAVEKEP----AR-----APPEPHNLFCAFGS 180

Db 184 TK--NPSCKMEELEFSVDSPYDVTGMMGARPGPDEGLSGLAPTPGPSSCKS- 239

Qy 181 WENELDHEMIEL 194

Db 240 -DGEGLDHWEEILV 252

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:58:11 ; Search time 40 Seconds
(without alignments)

324.958 Million cell updates/sec

Title: US-10-069-386-2

Perfect score: 1033

Sequence: 1 MEGGLKMRKHSDDLEEEERNE.....APGSWEWNELDHIMEIILGS 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters:

478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.HX:*

1: /cgn2_6_ptodata/1/iaa/5A_COMB.pep:*

2: /cgn2_6_ptodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6_ptodata/1/iaa/6A_COMB.pep:*

4: /cgn2_6_ptodata/1/iaa/6B_COMB.pep:*

5: /cgn2_6_ptodata/1/iaa/6C_COMB.pep:*

6: /cgn2_6_ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	9.1	433	4 US-08-259-451-3	Sequence 3, Appli
2	93	9.0	433	1 US-07-672-683-4	Sequence 4, Appli
3	93	9.0	2618	3 US-09-413-814-28	Sequence 28, Appli
4	92.5	9.0	432	5 PCT-0595-01910-13	Sequence 13, Appli
5	90.5	8.8	370	4 US-09-377-885B-16	Sequence 16, Appli
6	90.5	8.8	947	2 US-08-887-518-2	Sequence 2, Appli
7	90.5	8.8	947	2 US-09-023-321-2	Sequence 2, Appli
8	90.5	8.8	947	2 US-09-032-475-2	Sequence 2, Appli
9	90.5	8.8	947	3 US-09-257-103-1	Sequence 1, Appli
10	90.5	8.8	947	4 US-09-871-889A-1	Sequence 1, Appli
11	89.5	8.7	215	4 US-08-78-717-9	Sequence 9, Appli
12	88.5	8.6	897	4 US-09-849-002-18	Sequence 280, Appli
13	88	8.5	1130	4 US-09-976-994-280	Sequence 280, Appli
14	87.5	8.5	709	4 US-10-118-328-4	Sequence 4, Appli
15	86.5	8.4	136	4 US-08-59-451-5	Sequence 5, Appli
16	84	8.1	337	4 US-09-543-881A-7444	Sequence 7444, Appli
17	84	8.1	514	4 US-09-252-991A-25281	Sequence 25281, Appli
18	84	8.1	628	4 US-09-145-413E-48	Sequence 48, Appli
19	84	8.1	2152	3 US-09-036-987A-3	Sequence 3, Appli
20	84	8.1	2152	3 US-09-370-700-3	Sequence 3, Appli
21	84	8.1	2152	4 US-09-03-207-3	Sequence 3, Appli
22	83	8.0	969	2 US-08-148-159-1	Sequence 1, Appli
23	83	8.0	986	2 US-08-548-159-3	Sequence 3, Appli
24	83	8.0	1012	4 US-08-811-481-16	Sequence 16, Appli
25	83	8.0	1012	4 US-09-17-557-16	Sequence 16, Appli
26	82.5	8.0	1130	4 US-09-138-092-834	Sequence 834, Appli
27	82.5	8.0	2142	4 US-09-538-092-1142	Sequence 1142, Appli

ALIGNMENTS

28	82	7.9	667	4 US-09-248-796A-18663	Sequence 18663, Appli
29	82	7.9	1006	4 US-09-023-905A-12	Sequence 4, Appli
30	82	7.9	2101	1 US-08-466-390-4	Sequence 4, Appli
31	82	7.9	2101	1 US-08-470-950-4	Sequence 4, Appli
32	82	7.9	2101	1 US-08-467-781-4	Sequence 4, Appli
33	82	7.9	2101	1 US-08-195-487-4	Sequence 4, Appli
34	82	7.9	2101	3 US-08-483-924-2	Sequence 1, Appli
35	82	7.9	2101	3 US-09-452-294-1	Sequence 1, Appli
36	82	7.9	1597	3 US-09-423-890-13	Sequence 13, Appli
37	81.5	7.9	1597	3 US-08-628-829-14	Sequence 14, Appli
38	81.5	7.9	1597	3 US-08-252-991A-6979	Sequence 2, Appli
39	81	7.8	535	3 US-08-813-774-2	Sequence 2, Appli
40	81	7.8	513	4 US-09-732-025-2	Sequence 2, Appli
41	80.5	7.8	583	4 US-09-252-991A-32735	Sequence 46750, Appli
42	80.5	7.8	804	4 US-09-270-167-16750	Sequence 2, Appli
43	80.5	7.8	1162	2 US-08-728-323A-2	Sequence 2, Appli
44	80.5	7.8	1162	3 US-08-298-568-2	Sequence 2, Appli
45	80.5	7.8	1162	3 US-08-298-568-2	Sequence 2, Appli

RESULT 1 US-08-259-451-3	Sequence 3, Application US/08259451	Patent No. 6,006841	GENERAL INFORMATION:
			APPLICANT: Lee, Helen H.
			ASSIGNEE: Swanson, Priscilla A.
			ATTORNEY: Idler, Kenneth B.
			BUYER: Rosenblatt, Joseph D.
			CHIEF INVENTOR: Chen, Irvin S. Y.
			CO-APPLICANT: Goide, David W.
			CO-APPLICANT: Robertson, Eugene F.
			CO-APPLICANT: Stebbins, John E.
			CO-APPLICANT: Chan, Emerson W.
			CO-APPLICANT: Buyendorp, Mark H.
			CO-APPLICANT: Johnson, Joan E.
			CO-APPLICANT: Motley, Cheryl T.
			CO-APPLICANT: Peterson, Bryan
			CO-APPLICANT: Edwards, Michelle
			CO-APPLICANT: Guidinger, Peggy
			CO-APPLICANT: Tare, Cynthia
			CO-APPLICANT: Tare, Cynthia
			TITLE OF INVENTION: HTLV-1INRA Compositions for Detecting HTLV Infection
			NUMBER OF SEQUENCES: 19
			CORRESPONDENCE ADDRESS:
			ADDRESSEE: Abbott Laboratories
			STREET: One Abbott Park Road
			CITY: Abbott Park
			STATE: Illinois
			COUNTRY: USA
			ZIP: 60064
			COMPACT DISK FORM:
			MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
			MEDIUM TYPE: storage
			COMPUTER: IBM
			OPERATING SYSTEM: MS-DOS
			SOFTWARE: WordPerfect 5.1
			CURRENT APPLICATION DATA:
			APPLICATION NUMBER: US/08/259,451
			FILING DATE: 20-JUN-1994
			CLASSIFICATION: 426
			PRIOR APPLICATION DATA:
			APPLICATION NUMBER: US 08/086,415
			FILING DATE: 01-JUL-1993
			ATTORNEY/AGENT INFORMATION:
			NAME: Daniel W. Collins
			REGISTRATION NUMBER: 31,912
			REFERENCE/DOCKET NUMBER: 5381.US.P1
			TELECOMMUNICATION INFORMATION:
			TELEPHONE: (708) 937-6765

APPLICANT: States of America as represented by the Secretary, Department of Health and Human Services

TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A NOVEL PRIMATE T-CELL LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS: ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04910

FILING DATE: 21-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US08/231,526

FILING DATE: 22-APR-1994

CURRENT APPLICATION DATA:

NAME: WILLIAM S. FETTER

REGISTRATION NUMBER: 26,728

REFERENCE/DOCKET NUMBER: 2026-4125PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 755-4800

TELEFAX: 4212792

TELEFAX: 4212792

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acids

STRANDEDNESS: unknown

TOPOLOGY: unknown

PCT-US95-04910-13

Query Match 9.0% Score 92.5; DB 5; Length 432; Best Local Similarity 21.5%; Pred. No. 0; 92; Mismatches 19; Indels 79; Gaps 7; Matches 40; Conservative

QY 47 PRAPRLRAYLHTLQQLQALRPLAPATA-----LPPPEFLGEEDEFDLSA 93

DB 13 PKAP--PRGLSTHWNfLQASYLPGSDFDFQQLRFLKALKfIwLPIDfSLA 69

QY 94 TI-----GSILRELDTMDGTEPNPVTPGLQNEVPPQDPVFEALSSRYLGD 144

DB 70 SLIPRGYPERTSEITINVLRNQASPTPPAPSLP-----EPANPPPP-----111

QY 145 SGDDFFPLDIDTSAVEKEPARAPPEP-----PHNLFCAPGSWEWNE 185

DB 112 -----QGPs-APPEPHTPPPYIEPPATHCLPfLHPGAPSAAHRPWNKD 154

QY 186 LDHME 191

DB 155 LQAfKQ 160

RESULT 5 US-09-377-285B-16

Sequence 16, Application US/09377285B

Patent No. 6720175

GENERAL INFORMATION:

APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE

APPLICANT: WORLEY, Paul

APPLICANT: TU, Jian

APPLICANT: XIAO, Bo

APPLICANT: LEARY, Daniel

APPLICANT: BENKEN, Jutta

APPLICANT: LANGHAN, Anthony

TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)

FILE REFERENCE: JHU1380-4

CURRENT APPLICATION NUMBER: US/09/377-285B

CURRENT FILING DATE: 1993-08-18

PRIOR APPLICATION NUMBER: US 60/138,426

PRIOR FILING DATE: 1999-06-10

PRIOR APPLICATION NUMBER: US 60/138,493

PRIOR FILING DATE: 1999-06-10

PRIOR APPLICATION NUMBER: US 60/138,494

PRIOR FILING DATE: 1999-06-10

PRIOR APPLICATION NUMBER: US 60/097,334

PRIOR FILING DATE: 1998-08-18

NUMBER OF SEQ ID NOS: 72

SOFTWARE: PatentIn version 3.0

SEQ ID NO 16

LENGTH: 370

TYPE: PRT

ORGANISM: Rattus norvegicus

US-09-377-285B-16

Query Match 8.8% Score 90.5; DB 4; Length 370; Best Local Similarity 23.5%; Pred. No. 1; 2; Mismatches 29; Indels 49; Gaps 8; Matches 43; Conservative

QY 33 LLRISLDKV-QRSLGFRAPSLLRHYLHNTLQQLQALRPLAPALPP-----EPLFLG 85

DB 102 MYNMILEKVARREIGTLA-----TVVRLEPSQRVYIPESLPLPTPYCRKPL---147

QY 86 EEDFLSLSATGSIStRELDTMDGT-----EPQONPVTPLGLQNEVPPQ-PDPVFLAL-----137

DB 148 -NFACLDDVGHVfKDLSTOLSRISRTGLTSRKSIKAATPASATLERSPRIPEPVQLPAVfD 205

QY 138 -----SSRYLGDDSGLDDFLDIDTSAVEKEPARAPPEP-----PHNLFCAPG 179

DB 206 GRLSAASSVSSLASSAGSABEGASfTfQSfKGQVAPATfPPPIAPfTfPPfPLPAfTfPLPP 265

QY 180 SWE 182

DB 266 PME 268

RESULT 6 US-08-887-518-2

Sequence 2, Application US/08887518

Patent No. 5843721

GENERAL INFORMATION:

APPLICANT: Rohe, Mike

APPLICANT: Wu, Lin

TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 BUSH STREET, SUITE 3200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/887,518

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: T97-008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 343-4341.
 TELEFAX: (415) 343-4342.
 INFORMATION FOR SEQ ID NO: 2 :
 SEQUENCE CHARACTERISTICS:
 LENGTH: 947 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-023-321-2

Query Match 8.8%; Score 90.5; DB 2; Length 947;
 Best Local Similarity 25.8%; Pred. No. 4;
 Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

Query 3 GGLKRSKSDLUEEEERWENSPAGLQSYCQALLRISLDKVRSLGPRAPSLSRRLHVLINTL 62
 Db 659 GGLK--SPRGLEYKEPRHPPNQANYHQ----TIAQPRELSRAGPRAETGRA 710

Query 63 QQLQALRLAPAPALPPE----PLFLGEEFDSLSATIGSILRELDTSMDGTTEPPQNPVT 117
 Db 711 PRQL----PPLPBPPEPKNSPPLTSKSE----SGMW2PLPLSSLEPAPRNSS 758

Query 118 PLGLQNEVPPQ----PDPVFLALSSRY----LGDSGLDDEFLDIDTSAVEKEPA 164
 Db 759 P-ERATVPQELQOLEILFLNSLQPPSLEEQEQLISCLSLSDDS--EKNPS 814

Query 165 RAPPEPHNLFCAPESW----EWN 184
 Db 815 KASQSSRDTLSSGVHSSQAEARSSWN 843

RESULT 8
 US-09-023-475-2

Sequence 2, Application US/09032475
 ; Patent No. 5854033

GENERAL INFORMATION:
 APPLICANT: Rothe, Mike
 ATTORNEY/AGENT INFORMATION:
 APPLICANT: WU, Lin
 TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 STREET: 268 BUSH STREET, SUITE 3200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,321
 FILING DATE:
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/06/887,518
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: OSMAN, RICHARD A
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: T97-008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 343-4341.
 INFORMATION FOR SEQ ID NO: 2 :
 SEQUENCE CHARACTERISTICS:
 LENGTH: 947 amino acids
 TYPE: amino acid

Query Match 8.8%; Score 90.5; DB 2; Length 947;
 Best Local Similarity 25.8%; Pred. No. 4;
 Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

US-09-023-475-2

CURRENT APPLICATION NUMBER: US/09/871,889A
 CURRENT FILING DATE: 2001-06-01
 PRIORITY NUMBER: 09/257,703
 PRIORITY FILING DATE: 1999-02-25
 PRIORITY APPLICATION NUMBER: 60/076,299
 PRIORITY FILING DATE: 1998-02-27
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 1
 LENGTH: 947
 TYPE: PRT
 ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
 US-09-871-889A-1

Query Match 8.8%; Score 90.5; DB 3; Length 947;
 Best Local Similarity 25.8%; Pred. No. 4;
 Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

Qy 3 GGLKRKHSIDLEEEERWNPSAGLQSYQALLRISLDKVQRSLGPPAPSLSRRHVLTHNTL 62
 Db 659 GGLK--SPWRGEYKEPRHPPQANHQ----TLHQPRELSPPGPPRAEETTGRA 710

Qy 63 QQLQALRLAPALPPQ----PLFGEEDFSLSATGSLIRELDTMDTEPPONPVT 117
 Db 711 PKLQ----PPLPBPPEPPNPKSPPLTSEE----SGMWEPLPSLEPAPARNSS 758

Qy 118 PLGLQNEYVPPQ----PDVFLFELALSSRY----LGDSGLDDFLDIDTSAVEKEPA 164
 Db 759 P-ERKATVPEQELQQLERFLFLNSLQPSLSEQEQLSCLSDLSDDSS--ERKNPS 814

Qy 165 RAPPEPHNLFCAGSW----EWN 184
 Db 815 KASQSSRDTLSSGVHSWSSQAEARSSSN 843

RESULT 9
 US-09-257-703-1
 Sequence 1, Application US/09257703
 Patent No. 626538
 GENERAL INFORMATION:
 APPLICANT: Greene, Warner C.
 APPLICANT: Lin, Xin
 APPLICANT: Gelezunas, Romas
 TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALPHA AND IL-1
 FILE REFERENCE: 30448.6-USU1
 CURRENT FILING DATE: 1999-02-25
 EARLIER FILING DATE: 1998-02-27
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 1
 LENGTH: 947
 TYPE: PRT
 ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
 US-09-257-703-1

Query Match 8.8%; Score 90.5; DB 3; Length 947;
 Best Local Similarity 25.8%; Pred. No. 4;
 Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

Qy 3 GGLKRKHSIDLEEEERWNPSAGLQSYQALLRISLDKVQRSLGPPAPSLSRRHVLTHNTL 62
 Db 659 GGLK--SPWRGEYKEPRHPPQANHQ----TLHQPRELSPPGPPRAEETTGRA 710

Qy 63 QQLQALRLAPALPPQ----PLFGEEDFSLSATGSLIRELDTMDTEPPONPVT 117
 Db 711 PKLQ----PPLPBPPEPPNPKSPPLTSEE----SGMWEPLPSLEPAPARNSS 758

Qy 118 PLGLQNEYVPPQ----PDVFLFELALSSRY----LGDSGLDDFLDIDTSAVEKEPA 164
 Db 759 P-ERKATVPEQELQQLERFLFLNSLQPSLSEQEQLSCLSDLSDDSS--ERKNPS 814

RESULT 10
 US-09-871-889A-1
 Sequence 1, Application US/09871889A
 Patent No. 6645718
 GENERAL INFORMATION:
 APPLICANT: Greene, Warner C.
 APPLICANT: Lin, Xin
 APPLICANT: Gelezunas, Romas
 TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALPH
 FILE REFERENCE: 30448.61USD1

CURRENT APPLICATION NUMBER: US/09/871,889A
 CURRENT FILING DATE: 2001-06-01
 PRIORITY NUMBER: 09/257,703
 PRIORITY FILING DATE: 1999-02-25
 PRIORITY APPLICATION NUMBER: 60/076,299
 PRIORITY FILING DATE: 1998-02-27
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 1
 LENGTH: 947
 TYPE: PRT
 ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
 US-09-871-889A-1

Query Match 8.8%; Score 90.5; DB 3; Length 947;
 Best Local Similarity 25.8%; Pred. No. 4;
 Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;

Qy 3 GGLKRKHSIDLEEEERWNPSAGLQSYQALLRISLDKVQRSLGPPAPSLSRRHVLTHNTL 62
 Db 659 GGLK--SPWRGEYKEPRHPPQANHQ----TLHQPRELSPPGPPRAEETTGRA 710

Qy 63 QQLQALRLAPALPPQ----PLFGEEDFSLSATGSLIRELDTMDTEPPONPVT 117
 Db 711 PKLQ----PPLPBPPEPPNPKSPPLTSEE----SGMWEPLPSLEPAPARNSS 758

Qy 118 PLGLQNEYVPPQ----PDVFLFELALSSRY----LGDSGLDDFLDIDTSAVEKEPA 164
 Db 759 P-ERKATVPEQELQQLERFLFLNSLQPSLSEQEQLSCLSDLSDDSS--ERKNPS 814

Qy 165 RAPPEPHNLFCAGSW----EWN 184
 Db 815 KASQSSRDTLSSGVHSWSSQAEARSSSN 843

RESULT 10
 US-09-871-889A-1
 Sequence 1, Application US/09871889A
 Patent No. 6645718
 GENERAL INFORMATION:
 APPLICANT: Greene, Warner C.
 APPLICANT: Lin, Xin
 APPLICANT: Gelezunas, Romas
 TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALPH
 FILE REFERENCE: 30448.61USD1

CURRENT APPLICATION NUMBER: US/08/778,717
 CURRENT FILING DATE: 12-DEC-1996
 PRIORITY APPLICATION NUMBER: JP 352225/1995
 PRIORITY FILING DATE: 28-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2084-031-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 215 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: RECOMBINANT
 PUBLICATION INFORMATION:
 AUTHORS: NOBUYUKI FUJII ET AL.
 TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED
 FROM SAID FUSED DNA SEQUENCE AND METHOD OF
 EXPRESSING SAID FUSED PROTEIN
 RELEVANT RESIDUES IN SEQ ID NO: 9: FROM 1 TO 215
 US-08-778-71 7.9

Query Match Score 89.5; DB 4; Length 215;
 Best Local Similarity 23.7%; Pred. No. 0.72; Indels 41; Gaps 6;
 Matches 41; Conservative 24; Mismatches 67; Indels 41; Gaps 6;

Qy 2 EGGKLKRKHSDLEERE-----ERMEWSPAQLSYQQALLRISLKVRSLG----- 46
 Db 31 EGRAQPKVSEVIEDLYNCATAKEMEACVSAITIEACGSSLVPRGSEFMQIHLGLSPTP 90
 Qy 47 -PRAPSRLRHVLHNTLQLQALRLAPAPA-----LPPBPFLGEEDFSL 92
 Db 91 IPKAP--EGLSTHWHNLNQAYRLQPRPSDFDFQQLRFLKLTKPNNPDTYSL 147
 Qy 93 ATI-----GSIIRLEDTMDGTEPPONPVPVPLG---LQNEVPPQDPDPVLEA 136
 Db 148 ASLIPKGYGRVVEILINILVKNQVSPSAPAAVPPTCPTTTPPPPPSPSEA 200

RESULT 12
 US-09-849-602-18
 Sequence 18, Application US/09849602
 Patent No. 6794501
 GENERAL INFORMATION:
 APPLICANT: Scanian, Matthew J.
 APPLICANT: Old, Lloyd J.
 APPLICANT: Stockert, Elisabeth
 APPLICANT: Chen, Yao-Tsang
 TITLE OF INVENTION: Colon Cancer Antigen Panel
 FILE REFERENCE: L0461/7105 (JPN)
 CURRENT FILING DATE: 2001-05-04
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 18
 LENGTH: 897
 TYPE: PRT
 ORGANISM: Homo sapiens

US-09-849-602-18
 Query Match Score 88.5; DB 4; Length 897;
 Best Local Similarity 24.4%; Pred. No. 5.7; Indels 9; Gaps 9;
 Matches 50; Conservative 23; Mismatches 75; Indels 57; Gaps 9;

Qy 13 EEEPERWENSPAGLQSYQO-----AUJRSUDKVRSGRGRAP---SL 52
 Db 208 QQAER---QALQSLRQGGITLGKEMSTSS1PGCLGVALE---GDGSPHGHASL 256
 Qy 53 RRHVLHNTLQLQALRLAPAPLPEPLFGEEDFSLSATIGSILRELDTSMDGTEPP 112
 Db 257 LQHLL---LQEQRQSLTIAVLGGSPLVIGERVATSRVTKPRHRLSRTQSSP 312
 Qy 113 QNPVTPLQIYNEVPPQDPVFEALSSR---YLGDSGIDDDFLIDTSAVEKEPARAPPEP 170
 Db 313 -LPOSQPAQQLWQOQOFLKOKOOLQIGK----ILTKRGEELPRQPTTHEPEP 364
 Qy 171 PHNIFCAGSWEKNEHLMIEIILIG 195

RESULT 13
 US-09-976-594-280
 Sequence 280, Application US/09976594
 Patent No. 6673549
 GENERAL INFORMATION:
 APPLICANT: Furness, Michael
 APPLICANT: Buchbinder, Jenny
 PUBLICATION INFORMATION:
 TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 FILE REFERENCE: PA-0041-US
 CURRENT APPLICATION NUMBER: US/09/976,594
 CURRENT FILING DATE: 2001-10-12
 PRIORITY APPLICATION NUMBER: 60/240,409
 PRIORITY FILING DATE: 2000-10-12
 NUMBER OF SEQ ID NOS: 1143
 SOFTWARE: PERL Program
 SEQ ID NO 280
 LENGTH: 1130
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. 6673549 2258794CD1
 US-09-976-594-280

Query Match Score 8.5%; DB 4; Length 1130;
 Best Local Similarity 21.8%; Pred. No. 9.6;
 Matches 52; Conservative 23; Mismatches 75; Indels 88; Gaps 11;

Qy 5 LKRKHSDLEEEERNEWSAGL-----QSYQQALLRISLSD----- 39
 Db 539 LKR---LKAKEEPRWEEAQQGFNKIWRREQEYKYLK-SLDHQAVNFKNDTKALRSSL 593
 Qy 40 -----KVRQSLGPAPSRLR-VLHINTLQQLORALIAP---APALPPE--- 80
 Db 594 NEIESVYDHEQEHSEGRSAPSSEPHLIFYEDRQILDAALISYYKROPAIQKDQG 653
 Qy 81 -----PLFLGEEDFSLSATIGSILRELDTSMDGTEPPONPVPLGLONEVPPQPD 130
 Db 654 THQLLHQVPSLFFPSQLDGAEEADEDRSPQGTTDSERKRPAPGHSSPPEEK 713
 Qy 131 PVFLEALSSRYLGDSGLDDFFLDITSAVEKEPARAPPEP-----NLFCARGSW 181
 Db 714 GAF-----GDA-----PATEQPP--LPPPAHKPLDDYSLPFANNW 749

RESULT 14
 US-10-118-328-4
 Sequence 4, Application US/10118328
 Patent No. 6773504
 GENERAL INFORMATION:
 APPLICANT: YAN, Churihua et al.
 PUBLICATION INFORMATION:
 TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
 PROTEINS, AND USES THEREOF
 FILE REFERENCE: CL001220
 CURRENT APPLICATION NUMBER: US/10/118,328
 CURRENT FILING DATE: 2002-04-09
 PRIORITY APPLICATION NUMBER: 60/1282,460
 PRIORITY FILING DATE: 2001-04-10
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: FastSBQ for Windows Version 4.0
 SEQ ID NO 4
 LENGTH: 709
 TYPE: PRT
 ORGANISM: Mus musculus
 US-10-118-328-4
 Query Match Score 8.5%; DB 4; Length 709;

Best Local Similarity 26.6%; Pred. No. 5.3;

Matches 46; Conservative 18; Mismatches 62; Indels 47; Gaps 9;
 Query 5 LKRGHSDLBBEEERNEWSPGAGSYQQALLRISLD-----KVIQSLG----- 46
 Database 440 LKGNLINPEKRRKEWE-TIARIQQLQCRQVSLSPREPPIALAQRQSEQSYRVS 498

Query 47 -----PRAPILRRLVHLIHTLQ-LQQAIRLAP -APALIPEPIFLGEEFDSLAT 94
 Database 499 VIEPPAASSPSSPRIRRISLTKRLSALSREKNSPGGSPEDPSSP-----TSSVS 550

Query 95 IG---SILRELDTSMDGTBEPQNPTPLG--LQNEVPPQDPVLEA 140
 Database 551 PGSPSSPRNREPPPGSPASPGSPQSSFLKSLTMD-PPGFWPVTLITPSSSR 602

RESULT 15

US-08-259-451-5

; Sequence 5, Application US/08259451

; GENERAL INFORMATION:

; APPLICANT: Lee, Helen H.

; APPLICANT: Swanson, Priscilla A.

; APPLICANT: Idler, Kenneth B.

; APPLICANT: Rosenblatt, Joseph D.

; APPLICANT: Chen, Irvin S. Y.

; APPLICANT: Golde, David W.

; APPLICANT: Robertson, Eugene F.

; APPLICANT: Stephens, John E.

; APPLICANT: Chan, Emerson W.

; APPLICANT: Buyendorp, Mark H.

; APPLICANT: Johnson, Joan E.

; APPLICANT: Motley, Cheryl T.

; APPLICANT: Peterson, Bryan T.

; APPLICANT: Edwards, Michelle

; APPLICANT: Guidinger, Peggy

; APPLICANT: Tate, Cynthia

; TITLE OF INVENTION: HTLV-TINRA Compositions

; and Assays for Detecting HTLV Infection

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories

; STREET: One Abbott Park Road

; CITY: Abbott Park

; STATE: Illinois

; ZIP: 60064

; COUNTRY: USA

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

; MEDIUM TYPE: storage

; COMPUTER: IBM

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/259,451

; FILING DATE: 20-JUN-1994

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/086,415

; FILING DATE: 01-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Daniel W. Collins

; REGISTRATION NUMBER: 31,912

; REFERENCE/DOCKET NUMBER: 5381.US.P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (708) 937-6365

; TELEFAX: (708) 938-2623

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 136 amino acids

; TYPE: amino acid

; STRANDEDNESS:

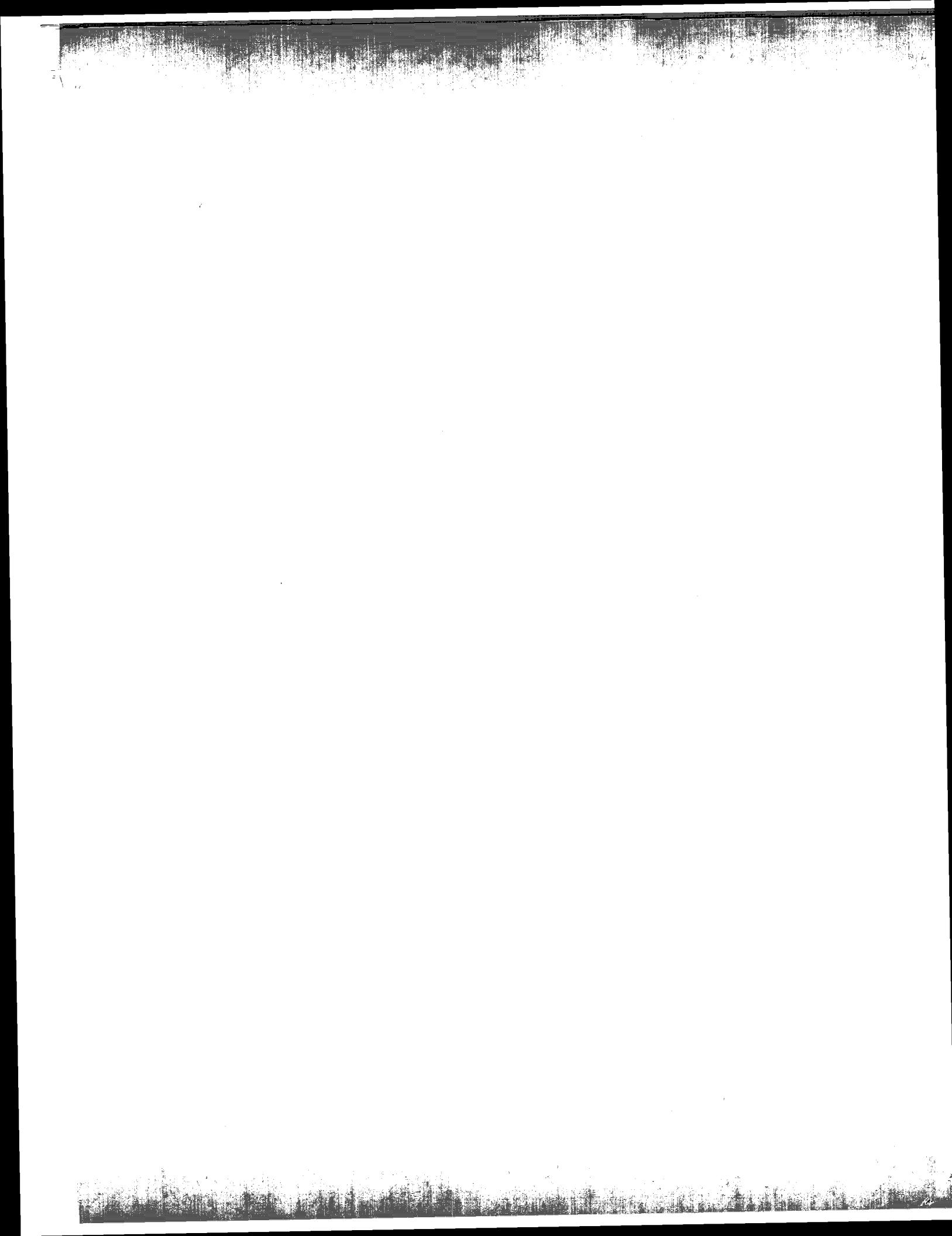
; TOPOLOGY: unknown

; US-08-259-451-5

Search completed: November 15, 2004, 14:08:14
 Job time: 43 secs

Query 47 PRAPSURRYLILHNTLQOQAAFLIAPAPA-----LPPPEPLFLGEEDESLSA 93
 Database 13 PKAP--RGISTHHTNFIQAAAYRQPGPSDFDFQQLRFLKALKTPWLNPTDYSLLA 69

Query 94 TI---GSILRELDTSMDGTBEPQNPTPLG--LQNEVPPQDPVLEA 136
 Database 70 SLIPKSYPPGRVVEINNLYQNQVSAPSAPAVPVTICPPTTTPPPPPPEPEA 121



Gencore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:57:50 ; Search time 38 Seconds

(without alignments)

496.276 Million cell updates/sec

Title: US-10-069-386-2

Perfect score: 1033

Sequence: 1 MEGGUKRKHSDLEEEERWE.....APGSWENNLDHIMEIILGS 196

Scoring table: BLOSUM62

GapOp 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : PIR_79;*

1: pir1;*

2: pir2;*

3: pir3;*

4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID Description

1	101.5	9.8	728	2	H59435	phosphoinositide-3
2	97	9.4	803	2	T40514	Chaperonin hsp78p
3	97	9.4	1201	2	G86441	unknown protein [i]
4	96	9.3	817	2	S51342	verrolin - yeast
5	95	9.2	446	2	A42029	transcription fact
6	93.5	9.1	1734	2	A54602	microtubule-associ
7	93	9.0	1433	1	FOLJH2	gag polyprotein -
8	91.5	8.9	1621	2	A82255	hypothetical prote
9	90.5	8.8	804	2	AG0565	probable membrane
10	88	8.5	651	2	T42644	hypothetical prote
11	88	8.5	716	2	T26998	hypothetical prote
12	88	8.5	1676	2	A56508	anucleate primary
13	86.5	8.4	536	2	A34596	transcription fact
14	86.5	8.4	1494	2	T14355	protein-tyrosine-p
15	86	8.3	475	2	T37761	ntA protein - Thi
16	85.5	8.3	485	2	T27550	hypothetical coile
17	85.5	8.3	825	2	E75508	conserved hypothet
18	85.5	8.3	1544	2	E59431	phosphoinositide-b
19	85	8.2	715	2	DB5087	hypothetical prote
20	85	8.2	1217	2	T00270	hypothetical prote
21	84.5	8.2	366	1	SAVLWD	large surface anti
22	84.5	8.2	477	2	T46304	hypothetical prote
23	84.5	8.2	723	2	B38749	3-phosphatidylinos
24	84.5	8.2	908	2	T16057	hypothetical prote
25	84.5	8.2	982	2	T43676	hunchback-related
26	84.5	8.2	1262	2	T13353	protein stn-B - fr
27	84	8.1	289	2	D70452	leucine-tRNA ligas
28	84	8.1	954	2	I61714	co-repressor prote
29	84	8.1	1015	2	JC5062	phogrin precursor

RESULT 2

T40514

Chaperonin hsp78p - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999

C;Accession: T40514

RESULT 1

T28968

hypothetical prote

hypothetical prote

hypothetical prote

probable AMP-bindi

probable oxidoredu

probable oxidoredu

hypothetical prote

DNA-binding prote

iichi protein - ink

structural protein

protein-tyrosine k

MHC class III hist

MHC class III hist

MHC class III hist

tailless (tl) pro

ALIGNMENTS

RESULT 1

H59435

phosphoinositide-3-kinase regulatory beta chain [imported]

- human

C;Species: Homo sapiens (man)

C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002

C;Accession: H59435; A59436

R;Vollini S; Patracchini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Roote L; V

Oncogene 7, 789-793, 1992

A;Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol

A;Reference number: A59435

A;Accession: H59435

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-728 <VOL>

A;Cross-references: GB:NP_005018; PID:g4826908; PID:NP_005018.1

A;Cross-references: GB:NP_005018; PID:g4826908; PID:NP_005018.1

Query Match 9.8%; Score 101.5; Pred. No. 1;

Best Local Similarity 23.9%; Mismatches 70; Indels 55; Gaps 9;

Matches 47; Conservative 25;

Qy 20 EMSPAD--GLOSYQAL-----

Db 164 QMDTAALADG-KSFLLALPAPLVTPPEASAAARRAAREAAAGPVGPALPEPTPLPHALTLR 223

Query 60 NTLQQLQAAIRLAPAPALPPEPLFLQEDDFLSLSATGSIURELDSDMGTEPPQNPVTPL 119

Db 224 FLIQLHGRVARRPA-----LGPAVTLGATFGPL-----LRAPEPPSSP-PPG 267

Qy 120 GLQNEVVPQPD--PVPEDEALSSRYLGDDFEIDTSAVERKEPARAPP----- 170

Db 268 GAPDSESPSPDPEALVKEQ-----EVAPPALPPRPPKARPAVTVLANGG 320

Qy 171 -PHNLFCAPGSEWNEEL 186

Db 321 SPPSL-QDAEWYWDI 335

R;Gwilliam, R.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M. Submitted to the EMBL Data Library, September 1998 Reference number: 22133	Db	439	ESSEEBHEFSH--HEHHEPAKOSVDN-PNLNLPDPPSSGDHVTL-----	DB 482
A;Accession: T40514	Qy	73	PAPALPPEPLIGEEDFSLSATIGSILRELDTSMDGTERBQNPTVPLGL-----	QNE 124
A;Status: preliminary; translated from GB/EMBL/DDBJ	Db	483	PPPEPPPPPLPTSTTSFSPS-----	QSPPPPPPLPMSTTSFSQPP 526
A;Molecule type: DNA	Qy	125	VPPQPPVPELALSSRYLGSGLDDFLIDTSAVEKEPARA2PEP 171	
A;Cross-references: UNIPROT:074402; EMBL:AB031534; PIDN:CAA20737.1; GSPDB:GN00057; SPDB:G	Db	527	PPPEPPPLF-----	TSTTSFSPSQPPPP 551
A;Experimental source: strain 972h-; cosmic c4f6				
A;Genes: SPBC4F6.17c				
A;Map position: 2				
C;Superfamily: ATP-binding chain				
F:1-139-146;Region: nucleotide-binding motif A (P-loop)				
F:207-212;Region: nucleotide-binding motif B				
F:539-546;Region: nucleotide-binding motif A (P-loop)				
F:607-612;Region: nucleotide-binding motif B				
F:145;Binding site: ATP (Lys) #status predicted				
F:545/Binding site: ATP (Lys) #status predicted				
Query Match 9.4%; Score 97; DB 2; Length 803;				
Best Local Similarity 24.2%; Pred. No. 28; Mismatches 31; Indels 62; Gaps 11;				
Matches 53; Conservative 31; Nucleotides 73; Amino acids 803;				
Matches 53; Amino acids 803;				
Query 5 LKRKHSIDLEBEER----NEWSPAGLQSTYQQA----LIRISLDKVORSLGPR-----49	Qy			
Db 377 LESTKDLXKEQDKLUSAAEERKULLDSIKRAKTELEBQRRIELETRQE-GNYVARASBLQ 435	Db			
Qy 50 ----PSLREV-----LINTLQQLQALLRLLAPAPALPPEPLIGED--RS 90	Qy			
Db 436 YAIIFBLERLVPKKEEKTLLKEKKPQMVHDDVTSDDIAVYVSATGSIPTNLKRGERDKILN 495	Db			
Qy 91 LSATIGSILRELDTSMDGTEPPQNPNVPTPGLQNEVPPQDPVPLTEAASSRYPGDSG---146	Qy			
Db 496 MEOTIGKKLIGQDEALKATAADAVR-LSRGLONTNP-----LASFLFELGPTGVKT 546	Db			
Qy 147 -----LDDFELDITSAV-----EKRP-ARAPPEP 171	Qy			
Db 547 ALTKALAEFLFDTKAMIRFDMSFQEQKHTIALLIGSP 585	Db			
Query Match 3				
Unknown protein [imported] - Arabidopsis thaliana				
C;Species: Arabidopsis thaliana (mouse-ear cress)				
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004				
C;Accession: G86442				
B;Theologis, A.; Becker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso-Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Chung, M.K.; Hughes, B.; Huzar, L.				
Nature 408, 816-820, 2000				
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzali, P.; Rizzo, M.; Rooney, T.-P.; Rowley, D.; Sakano, H.				
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.				
A;Reference: A86141; PMID:1130712				
A;Status: preliminary				
A;Molecule type: DNA				
A;Residues: 1-1201 <STOP>				
A;Cross-references: UNIPROT:Q9C6S1; GB:AE005172; PIDN:911136725; GSPDB:G				
A;Genetics:				
A;Map position: 1				
Query Match 9.4%; Score 97; DB 2; Length 1201;				
Best Local Similarity 24.6%; Pred. No. 47; Mismatches 11; Indels 62; Gaps 6;				
Matches 41; Conservative 11; Nucleotides 53; Amino acids 1201;				
Query 13 EEEERBWRWNSPAGLQSYQQALLRISLDKVQRSLGPRAPSRLRHHVLIHTLQQLQAAIRLA 72	Qy			
Db 594 DKSERLGMKKAKAPLPSHV-----PPPPVPP-----VLSDDSKNQVPA 633	Db			
Qy 96 GSILRELDTSMDGTEPPQNPNV-----TPLGLQNEVP-----PQDPFVFLREALSSRYL 142	Qy			

Db 634 ASLGDVLPSSNLKPPSPVAAAPPLPTESAPSILPQQSVSTSPSPSPV-APTLSTV-- 690 Qy 143 GDSGLDDFLFDIDTSAVEKEPARAP-PEPPHN 171 Db - 691 -----TETESISKNPTKSPPPPP 708

RESULT 5

A42029 transcription factor E3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A42029

R:Roman, C.; Materi, A.G.; Cooper, C.; Artandi, S.; Blain, S.; Ward, D.C.; Calame, K.

Mol. Cell. Biol. 12, 817-827, 1992.

A;Title: mTFE3, an X-linked transcriptional activator containing basic helix-loop-helix

A;Reference: A42029; MUID:92123207; PMID:1732746

A;Accession: A42029

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-446 <ROM>

A;Cross-references: UNIPROT:Q644092; GB:S756573; NID:3243429; PIDN:AA21130.1; PID:9243440

A;Note: sequence extracted from NCBI backbone (NCBIN:76673; NCBI:76674)

Query Match 9.2% Score 95.5; DB 2; Length 446;

Matches 48; Conservative 24; Mismatches 61; Indels 43; Gaps 8;

Db 272 LQEKEQR----SKOLESRQR----SLEQANRSLQRQEL----ELQAQHG 311

Qy 72 APAPALPPEPLFGEEDFLSISATIGSISIRELDTSMDGEPQNVPVPLQJLNEVPPQDDP 131

Db 312 LP---VPPNPGLLSPTTSSVSDSLKP---EQLDIEEGRESTFPHVGGPAQAPPQPPA 366

Qy 132 VFLEAL---SSRYVLDGSLGLDDFLFDIDTSAVEKEPARAPPEP 170

Db 367 PPSDDLLDHFPSDHLGDG-DPHLGLDILMEEGNYGGLGGALSPLRAASDP 421

RESULT 6

A54402 microtubule-associated serine/threonine protein kinase MAST205 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C:Accession: A54402

R:Walden, P.D.; Cowan, N.J.

Mol. Cell. Biol. 13, 7205-7235, 1993

A;Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associated with a novel testis-specific kinase. A54402; MUID:94067123; PMID:8243979

A;Accession: A54402

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1734 <WAL>

A;Cross-references: UNIPROT:Q60592; GB:U02213; NID:9406057; PIDN:AA04312.1; PID:9406058

C;Keywords: ATP

F:451-728/Domain: protein kinase homology <KIN>

F:459-467/Region: protein kinase ATP-binding motif

Query Match 9.1% Score 93.5%; DB 2; Length 1734;

Matches 51; Conservative 23; Mismatches 71; Indels 49; Gaps 8;

Db 1308 GSQSFPTKUHLSPPLGRPSAEPSPSLKR-----VQSAEKLAALAAE 1357

Qy 80 EPLFGEEDFLSISATIGSIRELDTSMDGEPQNVPVPLQNEVP---PQDPVFLEA 136

Db 1358 KKL-KPSRKHSIDLDPHGELKEL----TPREASPLEVVGTRSVLSSKGPLPGKGVLQP 1410

RESULT 7

FOLJH2 gag polyprotein - human T-cell lymphotropic virus type 2

N;Alternative names: core polyprotein

N;Contains: core protein p12; core protein p15; core protein p24

C;Species: human T-cell lymphotropic virus type 2, HTLV-2

A;Note: host Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 09-Jul-2004

C;Accession: A01944

R:Shimotomio, K.; Takahashi, Y.; Shimizu, N.; Gojobori, T.; Golde, D.W.; Chen, I.S.Y.; I. Proc. Natl. Acad. Sci. U.S.A. 82, 3101-3105, 1985

A;Title: Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type 2

A;Reference number: A94042; MUID:85216449; PMID:2582407

A;Cross-references: UNIPROT:PO3346; GB:MM10060; NID:9329559; PIDN:AA09884.1; PID:9329561

A;Note: the authors translated the codon TAC for residues 197 and 249 as Thr

A;Residues: 1-43 <SH1>

A;Molecule type: DNA

A;Accession: A03944

A;Genetics:

C;Superfamily: mammalian retrovirus gag polyprotein II

C;Keywords: core protein; Polyprotein

F:1-116/Product: core protein p15 #status predicted <P15>

F:137-214/Product: core protein p24 #status predicted <P24>

F:215-433/Product: core protein p12 #status predicted <P12>

Query Match 9.0% Score 93; DB 1; Length 433;

Best Local Similarity 23.1%; Pred. No. 2.7;

Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

Db 13 PKAP--RGLSTHNLNFQAYRQPRPSDFDFQQLRFLKAKTPILNPYDSLLA 69

Qy 94 TI----GSTLIRELDTSMDGEPQNVPVPLG--LQNEVPPQDPVFLEA-LSRVLG 143

Db 70 SLIPKPYGRVEIINLVNQVSISAPAPVPTICPTTTPPPPPSPEAHVPPY-- 127

Qy 144 DSGLDDFLFDIDTSAVEKEPARAPPEPFLCAGSWEENLDHIMEITLGS 196

Db 128 -----VEPTTQCFPLHPPGAP---SAHRPWNQMDLQAIKQEYSS 166

RESULT 8

A82255 hypothetical protein VC0998 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: A82255

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolava, M.D.; VanAken, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82255; MUID:0406333; PMID:10952101

A;Accession: A82255

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1621 <HEI>

A;Cross-references: UNIPROT:Q9KTA5; GB:AE004181; NID:99655454; PIDN:AAF9415

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0998

A;Molecule type: DNA
 A;Residues: 1-1676 <FIS>
 A;Cross-references: UNIPROT:Q000083 ; GB:X82289 ; NID:9683499 ; PMID:9683500
 C;Genetics:
 A;Gene: apsa
 A;Introns: 149/3

Query Match 8.5%; Score 88; DB 2; Length 1676;
 Best Local Similarity 20.5%; Pred. No. 42;
 Matches 32; Conservative 19; Mismatches 51; Indels 54; Gaps 6;

Qy 36 ISLDKVRQSPRAPSRLRHLVLIHNTLQLQALRLLAPAPALPPEPLIGEEDFLSLSATI 95
 Db 837 VSFEETE-PVAPSPELRTTAPFVGSTTPEVAPVPEVAVSP-----I 880

Qy 96 GSILRELDTSMDGTTRPQQVTPGLQNEVPPDDPVEI-----I 134
 Db 881 SS-----QRTQPTEVIP-----APPEDEPIVPPENAFSQLIVEDTLPILAXLP 924

Qy 135 FAISSLRYLGDSGLDDFLDIDTSAVEKEPARAPPEP 170
 Db 925 EPAPPRAFA-CGTTSDIADSVSASSEOTE-PEP 959

RESULT 13
 A34596 Transcription factor E3 - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 06-Jul-1990 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
 R;Beckmann, H.; Su, L.K.; Kadesch, T.
 Genes Dev. 4, 167-179, 1990
 A;Title: TFE3: a helix-loop-helix protein that activates transcription through the immun
 A;Reference number: A34596 ; MUID:0249724; PMID:2338343
 A;Accession: A34596
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-536 <EC>
 A;Cross-references: UNIPROT:P19532; EMBL:X51330; NID:937061; PID:91353
 A;Note: the authors translated the codon ACC for residue 433 as Ser, and GAG for residue
 C;Genetics:
 A;Gene: GDB:TFB3
 A;Cross-references: GDB:125970; OMIM:314310
 A;Map position: Xp11.23-Xp11.22
 C;Position: DNA binding; transcription factor

Query Match 8.4%; Score 86.5%; DB 2; Length 536;
 Best Local Similarity 25.4%; Pred. No. 13;
 Matches 47; Conservative 24; Mismatches 55; Indels 59; Gaps 10;

Qy 112 LEEEEBRWEMSPAGI-QSYQQALLRISLDKVQRSLGPRAPSLLRHLVLIINTLQQLQAAIRL 71
 Db 192 LQKEQR---SKLEDSQR---SIEEANRSLQIRQEL-----EIQAQH 231

Qy 72 APAPALPPEPLFLGFEDFLSATIGSISRE-LDTSMGSTE-----PPONPVTRPLG 120
 Db 232 LPVPGTP-----GLLSLATSTSDSLKPQLDIEEGSPGARTFHVGCGPAON-----279

Qy 121 LQNEVPPQP---DPVFLAEISSLRIGDSGLDDFLDIDTSAVEKEB-----PAR 165
 Db 280 APHQQPPAPPSDALLDLHFPSDHLDLG-DPFHIGLEDTILMEBEGVYGLSGALSSDLR 338

Qy 166 APPEP 170
 Db 339 AASDP 343

RESULT 14
 T14355 Protein-tyrosine-phosphatase (EC 3.1.3.48) TD14 - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C;Accession: T14355

R;Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M.
 J. Biol. Chem. 273, 21077-21083, 1998
 A;Title: A novel putative protein-tyrosine phosphatase contains a BROI-like domain and a
 C;Genetics:
 A;Reference number: 218004 ; MUD:98361981 ; PMID:694860
 A;Accession: T14355
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-194 <CDS>
 A;Cross-references: UNIPROT:OB8902; EMBL:AF077000; NID:93598973; PID:93598974; PIDN: AAC6
 A;Experimental source: brain
 C;Genetics:
 A;Gene: PTP-TD14
 C;Function:
 A;Description: may be involved in regulating Ha-ras-dependent cell growth
 C;Keywords: phosphotyrosine monoester hydrolase

Query Match 8.4%; Score 86.5%; DB 2; Length 1494;
 Best Local Similarity 29.1%; Pred. No. 48;
 Matches 37; Conservative 19; Mismatches 40; Indels 31; Gaps 9;

Qy 62 LQQLQAAIRLAPALPAPLPPPEPLGEGEDFLSATIGSILRE-----LDTSMGDTBPQNP 115
 Db 625 LMQRPAVPMAPGVPVLYPAPVYTS-----LGIVPRSPQHGVSSPYAGVGPQ-P 675

Qy 116 VTPGLQNEVPPQ-----PDPVFLEALSSRYLGDSGLDFFLDIDT-SAVEKEPARAPPEP 171
 Db 676 I-VGLFSAPPQFQSGP-----LAMDVR-PATTVDSVQAPISSMALRPGPAPRPPQ-P 727

Qy 172 HNLFCAP 178
 Db 728 ---CFP 730

RESULT 15
 B37761 ntrA protein - Thiohacillus Ferrooxidans
 C;Species: Thiohacillus Ferrooxidans
 C;Accession: B37761
 C;Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 09-Jul-2004
 R;Berger, D.K.; Woods, D.R.; Rawlings, D.E.
 A;Title: Complementation of Escherichia coli sigma(54) (ntrA)-dependent formate hydrogen
 A;Reference number: A37761 ; MUID:90330545; PMID:2198257
 A;Accession: B37761
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-475 <BER>
 A;Cross-references: UNIPROT:P24695; GB:M58480; GB:M33831; NID:9154642; PIDN:AAA27379.1;
 C;Superfamily: Pseudomonas transcription initiation factor sigma
 C;Keywords: DNA binding; transcription regulation

Query Match 8.3%; Score 86; DB 2; Length 475;
 Best Local Similarity 21.7%; Pred. No. 12;
 Matches 40; Conservative 31; Mismatches 75; Indels 38; Gaps 6;

Qy 1 MEGGLKRKHSDLEEEERWERSPAGLQSYQQALLRISLDKVQRSLGPRAPSLLRHLVLIINTLQQLQAAIRL 71
 Db 307 MAGGKDAAHKTIQDQLINEARFKISLQSRODTLKVARATVERQCDFFANGPESURPMVL 366

Qy 54 RHLVLIANTLQQLQAAIRLAPALPAPLPPFLGEEDFLSATIGSILRE-----SLGPRA--PSLR 53
 Db 367 RH-IADAVENHESTYSRVNTQKYMTRPGLYEFKYYFFSSVGTDSGGSSASATARRALLI 424

Qy 105 SMDGTEPPQNPVTPGLQNEVPPQDPVFILEALSSRYLGDSGLDDFLDIDTSAVEKEPA 164
 Db 425 KMTQAEAQHPLS-----DAEIRAVLADQGIQ---IARRTVARYREAA 464

Qy 165 RAPP 168
 Db 465 NVPP 468

Search completed: November 15, 2004, 14:07:21

Mon Nov 15 17:33:07 2004

us-10-069-386-2.rpr

Page 6

Job time : 40 secs